

Institute for Molecular Medicine Finland, FIMM
University of Helsinki, Finland
The Doctoral Programme in Biomedicine (DPBM)

CLONAL EVOLUTION AND HETEROGENEITY OF CANCER IN THE CONTEXT OF INDIVIDUALIZED MEDICINE

Poojitha Ojamies

ACADEMIC DISSERTATION

To be presented, with the permission of the Faculty of Medicine, University of Helsinki, for public examination in Haartman-Institute, Lecture Hall 2, Haartmaninkatu 3, Helsinki on 24th of August 2018, at 12 o'clock noon.

Helsinki, 2018

Supervised by	Professor Olli Kallioniemi, M.D., Ph.D Institute for Molecular Medicine Finland HiLIFE, University of Helsinki, Finland
	Docent Maija Wolf, Ph.D. Institute for Molecular Medicine Finland University of Helsinki, Finland
Thesis Committee	Docent Petri Bono M.D., Ph.D Helsinki University Hospital University of Helsinki, Finland
	Harri Lähdesmäki, D.Sc. (Tech) Aalto University School of Science, Finland
	Docent Dario Greco, Ph.D University of Tampere, Finland
Reviewed by	Docent Olli Lohi, M.D., Ph.D. Pediatric Hematology & Oncology Division Tampere University Hospital, Finland
	Docent Merja Heinäniemi, Ph.D, Dos. Institute of Biomedicine University of Eastern Finland, Finland
Official opponent	Dr Nicholas McGranahan, Ph.D. Department of Oncology University College London, UK

ISBN 978-951-51-4403-4 (paperback)
ISBN 978-951-51-4404-1 (PDF)

*Dedicated to my beloved grandparents,
Kathyayini and Venkateswarlu.*

TABLE OF CONTENTS

LIST OF ORIGINAL PUBLICATIONS.....7

ABBREVIATIONS.....8

ABSTRACT9

INTRODUCTION 11

REVIEW OF THE LITERATURE 14

1. Background 14

 1.1 Mutational landscape of cancer 14

 1.1.1 Genomic alterations in AML..... 15

 1.1.2 Genomic alterations in RCC..... 16

2. Cancer evolution and heterogeneity 18

 2.1 A brief history..... 18

 2.2 Evolution in cancer 19

 2.2.1 Clonal origin of cancer..... 20

 2.2.2 Clonal evolutionary trajectories 22

 2.3 Intra-tumor heterogeneity of cancer..... 24

3. Clinical impact of tumor evolution and heterogeneity 26

 3.1 Impact of therapy on clonal evolution 26

 3.2 Emergence of drug resistant subclones 27

 3.2.1 Clonal and subclonal mutations..... 27

 4. Longitudinal and spatial sampling of tumors 28

AIMS OF THE STUDY..... 29

MATERIALS AND METHODS..... 30

1. Patient study material - AML..... 30

2. Study Material - RCC 30

2.1 Patient samples	30
2.2 Patient derived cultures (PDC)	31
3. Genome sequencing technologies.....	31
3.1 Whole exome sequencing analysis	31
3.2 Ultra-deep amplicon sequencing	32
3.2.1 Unique molecular identifier (UMI) based ultra-deep amplicon sequencing.....	33
4. Drug sensitivity and resistance testing (DSRT).....	34
5. Statistical analysis	34
5.1 Correlation matrices	34
6. Phylogenetic analysis.....	35
RESULTS	36
1. Monitoring therapy response at subclone-level	36
1.1 Personalized medicine platform for analysing subclonal therapy response	36
1.2 Methodologies for interpreting subclonal architecture and therapy response from multiple time point samples	37
1.3 Subclonal responses to chemotherapy	39
1.4 Subclonal responses to targeted therapy	41
2. Genetic heterogeneity between spatially distinct bone marrow sites in AML	46
3. Genetic heterogeneity underlying drug responses in renal cell carcinoma	48
3.1 Genetic heterogeneity and relation between multiple tumor sites and established PDCs.....	49
3.2 Impact of genomic heterogeneity on drug response profiles	50
DISCUSSION	52

1. Subclonal responses to therapy.....	52
2. Mechanistic insights on genetic heterogeneity underlying drug response and resistance.....	54
3. Lessons learnt from personalized treatment of patients	56
4. Next Steps: targeting clonal heterogeneity and evolution.....	58
ACKNOWLEDGMENTS	62
REFERENCES	65

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following publications:

- I. **Poojitha N. Ojamies**, Mika Kontro, Henrik Edgren, Pekka Ellonen, Sonja Lagström, Henrikki Almusa, Timo Miettinen, Samuli Eldfors, David Tamborero, Krister Wennerberg, Caroline Heckman, Kimmo Porkka, Maija Wolf, Olli Kallioniemi.
Monitoring therapy responses at the leukemic subclone level by ultra-deep amplicon resequencing in acute myeloid leukemia. *Leukemia*. 2017 May;31(5):1048.
- II. **Poojitha N. Ojamies**, Mika Kontro, Henrik Edgren, Pekka Ellonen, Caroline Heckman, Kimmo Porkka, Maija Wolf* and Olli Kallioniemi*. Case studies investigating genetic heterogeneity between anatomically distinct bone marrow compartments in acute myeloid leukemia. *Leukemia & lymphoma*. 2018 Apr 5:1-4.
- III. Khalid Saeed, **Poojitha N. Ojamies**, Teijo Pellinen, Samuli Eldfors, Riku Turkki, Johan Lundin, Harry Nisen, Taija M. af Hällström, Antti Rannikko, Tuomas Mirtti, Olli Kallioniemi, Päivi Östling.
Clonal heterogeneity influences drug responsiveness in renal cancer - assessed by *ex vivo* drug testing of multiple patient-derived cancer cells. Manuscript Submitted

* The authors contributed equally to this work.

The articles are reproduced with permission from the copyright holders.

ABBREVIATIONS

AML	Acute myeloid leukemia
BM	Bone marrow
BP	Base pair
ccRCC	Clear cell renal cell carcinomas
CML	Chronic myelogenous leukemia
CNV	Copy number variations
CRC	Conditionally reprogrammed cells
CS	Corpus sternum
CTG	CellTiter-Glo
DSRT	Drug sensitivity and resistance testing
DSS	Drug sensitivity score
ELN	European LeukemiaNet
HSPC	Hematopoietic stem progenitor cells
HUB	Helsinki Urological Biobank
ISM	Individualized systems medicine
ITH	Intra-tumoral heterogeneity
LIC	Left iliac crest
MDS	Myelodysplastic syndrome
PB	Peripheral blood
PCR	Polymerase chain reaction
PDCs	Patient-derived cell cultures
RCC	Renal cell carcinoma
RIC	Right iliac crest
RPKM	Reads per kilobase of target region length per million mapped reads
TCGA	The Cancer Genome Atlas
UMI	Unique molecular identifier
VAF	Variant allele frequencies
WES	Whole exome sequencing
WHO	World Health Organization

ABSTRACT

Cancer is an evolutionary phenomenon. Development and progression of cancer involves genetic instability and a complex clonal evolution process, which contributes to genetic and phenotypic heterogeneity. Advent of next-generation sequencing technologies has enabled in-depth interrogation of cancer genomes resulting in unraveling of the complexity of genomic heterogeneity and evolution in detail. Studies on relapsed and metastatic patients have revealed that subclones present at low-frequency at diagnosis could lead to relapse or seeding of distant metastatic clones, underscoring the importance of detecting rare subclones early on.

The dynamic nature of clonal evolution and heterogeneity further adds to the complexity of treating cancers and difficulty in achieving long-term remissions. Tumor heterogeneity should be taken into account in developing new therapeutic strategies to target cancer at the subclone level and pre-emptively block the emergence of rare subclones that could lead to relapse. Therefore, as part of our individualized systems medicine program, we developed methods to interrogate spatial and temporal heterogeneity and evolution of cancer subclones to delineate the impact of therapy on them.

In study I, we monitored responses of chemo and targeted therapies at leukemic cell subclones in chemorefractory acute myeloid leukemia. We performed exome-sequencing analysis of serial samples acquired from diagnosis to later relapses from the 13 AML patients. In order to extend our understanding of clonal composition to include low-frequency subclones ($< 10\%$ frequency), we developed unique molecular identifier based ultra-deep amplicon sequencing approach to enhance rare variant detection. Statistical methods to interpret response of subclonal mutations to therapy from longitudinally acquired samples were employed to gain mechanistic insights on therapy response and resistance.

In study II, we interrogated if genetic heterogeneity exists between multiple bone marrow compartments in AML patients. Numerous recent studies have reported on the extent of intra-tumoral heterogeneity in solid tumors, however very little information is available on its role in AML. Therefore, we acquired biopsies of

three anatomically distinct bone marrow compartments: corpus sternum, right and left iliac crest from two AML patients. Based on phylogenetic analysis of mutations identified from exome-sequencing data we reconstructed the cancer cell lineage from early ancestral cell to later diversification of subclones to distinct bone marrow compartments in individual patients. Although we did not identify subclonal mutations specific to a bone marrow site, we observed significant variance in subclonal frequencies between the bone marrow compartments and could interpret the genetic interconnection between the compartments.

In study III, we integrated intra-tumor genomic heterogeneity with drug response heterogeneity in renal cell carcinoma to understand the impact of the cancer heterogeneity and evolution on variability in drug responsiveness. We developed patient derived cell culture models from primary and metastatic tumor tissues of four renal cell carcinoma patients for the study. Based on phylogenetic analysis of somatic mutations in tumor tissue and patient-derived cell culture models we inferred i) how representative the models are of the originating tumor tissues, ii) evolution and divergence of subclones in tumor tissues and models and, iii) the drug sensitivity and resistance profiles of each patient-derived cell culture model and variability of drug response between individual subclones.

Taken together, we have developed methods to comprehensively study subclonal heterogeneity, subclonal-level responses to chemo and targeted drugs and integrated genetic heterogeneity with drug response profiles in a personalized medicine setting. The methods showcased in our studies are applicable to all cancers and can be tailored for longitudinal and spatial sampling strategy of tumors during the disease course. We believe that understanding of subclonal evolution and heterogeneity will provide basis for devising better combinatorial strategies to prevent future relapses to achieve better remission rates by targeting multiple key subclones.

INTRODUCTION

Cancer is not a singular disease but a phenomenon of accelerated evolution that continually changes the cancer genotype. New mutational events lead to evolution of cancer cells under selective pressure of the host and tumor microenvironment or cancer treatment. Treatment of cancer with chemotherapeutics, especially genotoxic drugs like alkylating agents not just eradicate major cancer clones but unleash additional mutations and a heterogeneous drug resistant cancer cell population. As a consequence of the resulting heterogeneity, cancer cells undergo evolutionary selections leading to emergence of drug resistant subclones and subsequent re-emergence of cancer. To understand the evolutionary nature of cancer and the impact of various factors on clonal selection, it is pertinent to obtain and analyze sequential cancer samples from diagnosis to eventual disease progression and therapy.

The “concept of cure” in cancer was actualized in early 1960’s when significant remissions and long-term progression free survival were achieved using chemotherapeutics.¹ Chemotherapy has been very effective and achieved significant “cures” in terms of progression-free survival in a subset of cancers such as paediatric acute leukemias, Burkitt’s lymphoma and in solid tumors such as testicular cancer, non-small cell lung cancer and subtypes of breast cancers.¹⁻⁶ Although, the overall contribution of cytotoxic therapy to progression-free survival remained at a modest 2.1% in adult cancer malignancies.² In AML, chemotherapy is successful in achieving remissions in younger (60-80%) and older (60%) patients. Although 85% of older patients relapse within 2-3 years.^{7,8} Relapse in cancer can be caused due to drug resistance, heterogeneity and selective pressure induced by therapy resulting in clonal evolution.^{9,10}

In late 90’s, a new generation of ‘targeted therapies’ designed to specifically target key molecular mechanisms involved in tumor initiation or growth emerged. The aim was to design drugs with high specificity and low toxicity. Discovery of BCR-ABL inhibitor (Gleevec) for treatment of chronic myelogenous leukemia (CML) heralded the beginning of biomarker based drug discovery for the next few decades.¹¹ With the completion of human genome project and advent of new sequencing technologies, extensive molecular

characterization of tumors became possible and resulted in designing of many new targeted therapies. Although kinase inhibitors which constitute majority of the targeted inhibitors, showed effect in cancers driven by single mutations in defined kinase-pathways, the magnitude of response in other cancers has been more limited. Cancers initially responding to targeted inhibitors relapsed due to positive selection of intrinsic drug-resistant cell clones or emergence of drug resistant subclones as a consequence of therapy.^{9,12,13}

In most of the common hematological and solid tumors, targeted therapies have not resulted in a dramatic progress in long term remissions and cancer cures as was expected. This underscores the importance of the evolutionary aspect in cancer that heterogeneous tumors are hard to eradicate. Therefore, understanding of cancer evolutionary processes is required to make progress in managing cancer, if not for curing it. Recent advances in sequencing of sequential samples from hematological and solid tumors have shed light on cancer evolution and the underlying inter- and intra-tumor heterogeneity, thereby reshaping our thinking of cancer as even more complex and dynamic disease than earlier thought.

The difficulties in introducing ‘one drug for all patients’ concept have led to the consideration of treating cancer at a personalized level in the hope of achieving long-term remissions by designing intelligent drug regimens tailored for a given patient. Therefore, many consortiums including ours have put efforts into deep profiling of serial samples from individual patients to design drug combinations targeting multiple cell populations in the hope of achieving a better outcome for the patient.

In this thesis, we sought to expand our interpretation of clonal architecture of a tumor to include low-frequency subclones and developed methods to assess treatment response at subclone-level with acute myeloid leukemia (AML) as a model. We studied intra-tumoral heterogeneity (ITH) and its role in heterogeneous drug responses in AML and renal cancers. These studies provide methods to investigate cancer evolution and heterogeneity as well as understand drug responses at a subclone level. We believe, our personalized medicine-based approach to target cancer heterogeneity by designing smart drug combinations

targeting multiple subclones will probably reduce the evolutionary potential of a tumor by making it more homogenous and help reduce future relapses.

REVIEW OF THE LITERATURE

1. BACKGROUND

1.1 Mutational landscape of cancer

Before we dive into exploring clonal evolution, it is pertinent to gain a comprehensive understanding of the mutational landscape of cancer and underlying processes that lead to accumulation of cancer mutations over time. Seminal paper by Vogelstein et al., reviewing the genomic landscapes of multiple cancers highlighted the diversity in mutational load of different cancers, although majority of the driver alterations converged into 12 key signaling pathways.¹⁴ Tumors of the colon and lung as well as melanomas were among the top highly mutated tumor types with > 200 nonsynonymous mutations per tumor whereas leukemia's were among the least mutated with < 25 mutations per tumor. Large chromosomal alterations including deletions, translocations and inversions were found to be frequent in solid tumors such as colorectal and breast cancers. With such diverse genomic alterations underlying each tumor, it is important to differentiate and categorize mutations that confer survival advantage for a tumor i.e., driver versus passenger mutations.

Several cancer biology studies over time have identified mutations in genes such as *TP53*, *APC*, *KRAS* and their importance in driving cancers. Recent pan-cancer tumor sequencing efforts by consortia's such as The Cancer Genome Atlas (TCGA) have provided a comprehensive list of recurrent gene aberrations in >30 major cancer types, laying the groundwork for future cancer genomic investigations. Genes recurrently mutated in tumors indicate high likelihood of being crucial in driving the tumor. Although, for a mutation to be considered as a driving event, additional features such as functional significance of a mutation, whether it is in conserved regions, phosphorylation status, the variant allele frequency and the presence of multiple other mutations in the same gene need to be taken into consideration.¹⁵ Recent studies have published methods to identify such driving alterations from publicly available consortium data, thereby providing a comprehensive catalogue of driver genes in major cancers.¹⁵⁻²³

1.1.1 Genomic alterations in AML

AML is a heterogeneous disease characterized by evolving genomic landscape and acquisition of new mutations. Recent studies emerging from massive sequencing efforts have revealed comprehensive list of mutations driving AML from preleukemic phase to late relapse.²⁴⁻²⁶ In AML, the most common recurrent mutations occur in *FLT3*, *NPM1*, *DNMT3A*, *NRAS*, *TET2*, *IDH1/2* and *RUNX1* driver genes. Chromosome level alterations such as chromosome 7/7q(del), 8/8q(gain), 5/5q(del) are highly recurrent.²⁴ Based on World Health Organization (WHO) classification, molecular subgroups of AML were defined based on specific set of chromosomal alterations and translocations.²⁷ Re-evaluation of AML classification into subgroups based on cytogenetic analysis and sequencing of 111 key genes from 1540 AML patient samples identified 11 distinct classes of AML as shown in Figure 1. *NPM1*-mutated patients, with additional mutations in *DNMT3A*, *FLT3*, *NRAS*, *PTPN11* and *TET2* made up the largest cohort (27%). Interestingly, mutations in genes regulating RNA splicing (*SRSF2*, *SF3B1*, *ZRSR2*, *U2AF1*) and chromatin regulation mechanism (*ASXL1*, *STAG2*, *EZH2*, *PHF6*), which were previously not classified based on WHO, made up the second largest cohort (18%).²⁴ Recent studies have highlighted the incidence of age-related clonal hematopoiesis in non-AML healthy adults. Mutations in *DNMT3A*, *TET2* and *ASXL1* genes were found to be frequent (9-18%) in healthy individuals with >70yrs of age. These mutations are associated with increased risk of hematological and cardiovascular diseases.²⁸⁻³⁰

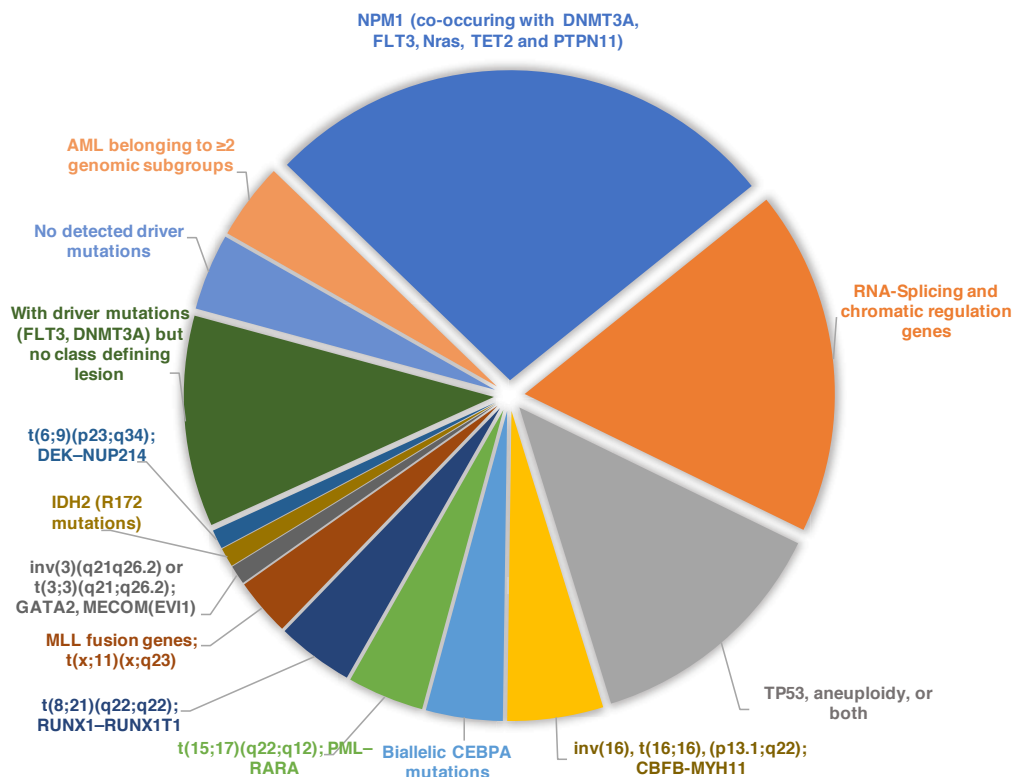


Figure 1: Genomic classification of AML into 11 major subgroups based on cytogenetic and sequencing analysis of 1540 AML patient samples. *NPM1* mutation with co-occurring alterations in *DNMT3A*, *TET2*, *FLT3*, *NRAS* and *PTPN11* is the most recurrent followed by alterations in RNA splicing and chromatic regulating genes (*RUNX1*, *MLL*, *SRSF2*, *DNMT3A*, *ASXL1*, *STAG2*, etc).

Reproduced with permission from “Genomic Classification and Prognosis in Acute Myeloid Leukemia” Papaemmanuil et al.²⁴ Copyright Massachusetts Medical Society 2016.

1.1.2 Genomic alterations in RCC

Clear cell renal cell carcinoma (ccRCC), the common type of cancer accounting to > 90% of kidney cancers, is often characterized by significant genomic heterogeneity at both inter and intra-tumoral level.³¹ Large-scale sequencing efforts have identified recurring somatic alterations in driver genes such as *VHL*, histone modification and chromatin remodeling genes (*PBRM1*, *SETD2*, *KDM5C*) and genes belonging to PI(3)K/AKT/MTOR pathway.³²⁻³⁵ *VHL*

mutations are highly recurrent in ccRCCs and seen in up to 90% of ccRCC cases. Its inactivation leads to degradation of hypoxia inducible factor (HIF α) transcription factor family leading to a range of downstream effects in pathways involved in angiogenesis, glycolysis and anabolic biosynthesis.³⁶⁻³⁹ Recurrent mutations in PI(3)K/AKT/MTOR pathway have a critical impact on cell growth and proliferation leading to tumor progression.⁴⁰ Therefore, MTOR inhibitors are highly interesting therapeutic agents for inhibiting PI(3)K/AKT pathway.^{41,42} Although, underlying intra-tumoral heterogeneity between distinct tumor regions of ccRCC's complicates the disease and its treatment.^{43,44}

2. CANCER EVOLUTION AND HETEROGENEITY

2.1 A brief history

‘Nothing in biology makes sense except in the light of evolution’

Theodosius Dobzhansky

Peter Nowell introduced the concept of cancer as an evolutionary phenomenon similar to the process of Darwinian evolution in 1976, nearly four decades ago.⁴⁵ But, not until recently has the concept of competing evolution among cancer clones been revisited.^{43,46-51} Pivotal studies on clonal architecture and evolution of somatic mutations in acute lymphoid and myeloid leukemia beautifully described the origin and evolution of subclones from diagnosis to later relapses, while capturing the impact of therapy on reshaping the tumor subpopulation.^{46,49,52} So what enabled us to re-establish cancer as an evolutionary mechanism which was first proposed over four decades ago?

The first human genome sequence was published after a decade of massive collaborative effort in 2001.⁵³ Since then, the price of sequencing has dramatically fallen, even exceeding the Moores law.⁵⁴ Reduction in sequencing costs led to a barrage of cancer genome sequencing efforts resulting in detailed evaluation and cataloguing of cancer genome and associated mutations. The coordinated acquisition of cancer genome data led to the identification of novel cancer driver events, reclassifying tumors based on their molecular profile, improved diagnostics and development of new targeted therapies. Another key concept of cancer that gained significant leverage with advancement of next-generation high throughput sequencing was the existence of genetic heterogeneity in tumors. Many recent studies on genome sequencing of tumors from multiple time points during the course of disease progression provide evidence of clonal evolution and selective pressure of therapy on cancer subclones in an unprecedented detail. Evolution of cancer, a concept introduced four decades ago, has now been demonstrated effectively with the advent of next-generation sequencing technologies enabled in-depth interrogation of the genome.

2.2 Evolution in cancer

The Darwinian theory of evolution by natural selection⁵⁵ is based on the following three basic principles:

1. Diverse group of organisms evolve from common shared ancestors
2. Phenotypic variations that are heritable are acquired in each generation
3. These variations confer natural selection, wherein positive traits are selected for the next generation resulting in “survival of the fittest”

This Darwinian model system of evolution is reminiscent of how cancers evolve overtime. A cancer cell is basically a normal cell gone rogue with unchecked proliferation and invasive capacity while accumulating mutations that confer survival advantage over normal cells. Cancer cells evade checkpoints such as programmed cell death, apoptosis and immune surveillance resulting in accumulation of abnormal cells leading to tumor formations. Additionally, tumors are able to recruit normal cells and hack the cellular functions of the body to build a sustainable tumor microenvironment that supports the growth of the tumor and protects it from adverse effects such as therapy.⁵⁶⁻⁵⁹ Such a complex system of disease, evolves over time and is constrained by a range of selective pressures such as therapy, environmental effects and microenvironment as shown in Figure 2.

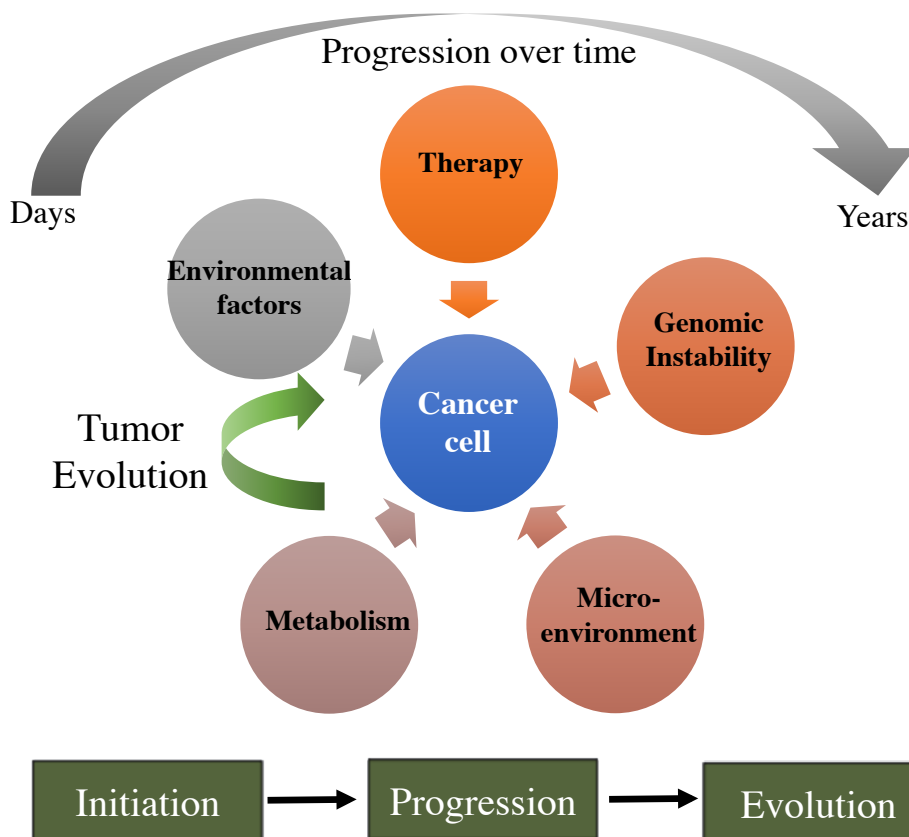


Figure 2: Factors influencing evolution of cancer cells from initiation to progression. Factors such as therapy and environment influence a cancer cell state resulting in genomic and epigenomic changes that lead to cancer initiation, progression and evolution over time.

Adapted with permission from Springer Nature Publisher: Nature Medicine, “Toward understanding and exploiting tumor heterogeneity” Alizadeh et al.⁶⁰ Copyright 2015

2.2.1 Clonal origin of cancer

Cancer usually originates at cellular level, wherein a progenitor or normal cell stochastically accumulates heritable alterations that undergo selective pressure resulting in ‘survival of the fittest clones’. Every cell in human acquires mutations at a constant rate while aging, but these are usually “passenger” variations that do not result in extensive disruption of the cellular functions.^{52,61,62} However, when cells acquire “driver” alterations in key cancer genes such

as oncogenes, tumor suppressors or DNA repair genes, this could lead to initiation of tumorigenesis. The number of driver alterations required to give rise to a malignant cancer cell varies between different cancers. Successive mutations acquired by the tumor initiating cells confer selective advantage to the cells leading to tumor progression. The availability of large-scale cancer genomic data has enabled the cataloguing of driver alterations in various cancers.^{15,18,63} The driving alterations could be predominantly point mutations or translocations in cancers such as AML, or copy number and structural alterations such as in prostate and ovarian cancers.^{25,52,64-66}

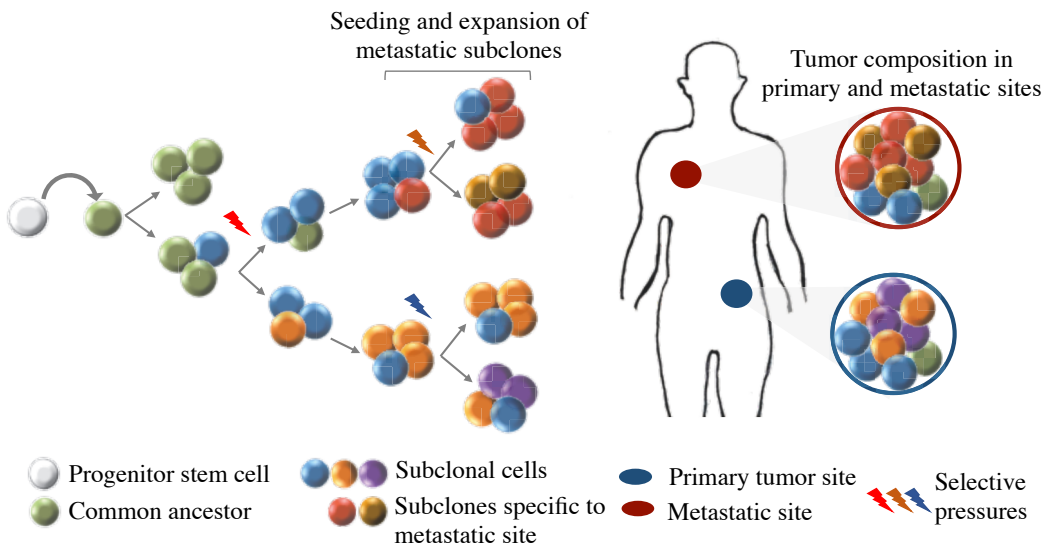


Figure 3: Clonal evolutionary model of cancer. Cancer clones arising from a common ancestor acquire additional mutations and branch out to subclones. Some of these subclones gain metastatic potential and seed tumor in a distant metastatic site. A snapshot of the primary and metastatic tumor shows heterogeneous population of cancer cells to include both early ancestral cell clones and subclones arising during the evolution of cancer. Adapted with permission from Springer Nature Publisher: Nature, “Clonal evolution in cancer” Mel Greaves, Carlo C. Maley,⁶⁷ copyright 2012.

Tumor initiating events happen at the cancer stem cell stage as seen in AML and CML wherein pre-leukemic mutations are acquired overtime by hematopoietic stem progenitor cells (HSPC).⁶⁸ Cooperating mutations accumulated in HPSCs form a founder clone present in all tumor cells. Subsequently, the tumor progresses by gaining additional mutations resulting in branching out of many

different cancer subclones (Figure 3).^{52,60,67-69} Selective pressures induced by factors such as therapy and chromosomal instability further shape the clonal evolutionary process. As a consequence of evolution, heterogeneous cancer subclones are formed resulting in intra-tumor heterogeneity. ITH in turn results in competition between multiple subclones leading to competing evolution thereby resulting in a circular problem of evolution leading to ITH and vice versa.

2.2.2 Clonal evolutionary trajectories

Cancers can evolve in a linear fashion, wherein mutations are acquired sequentially by the major clone resulting in subclones that can be traced back to the ancestral cancer clone. Alternatively, in branched evolution, different mutations are acquired by cancer cells resulting in diversification of the tumor into multiple subclones (Figure 4). Studies on clonal evolutionary trajectories have identified both linear (as shown in acute lymphoid leukemia⁴⁶) and branched trajectories from a recent common ancestor (as shown in primary breast cancers⁷⁰). These studies have shown how therapy can impact cancer evolution by creating evolutionary bottlenecks. Therefore, to understand the overall life-history of a cancer and how external and internal factors such as therapy impact the evolutionary structure we must obtain and analyze longitudinal samples during the disease course.

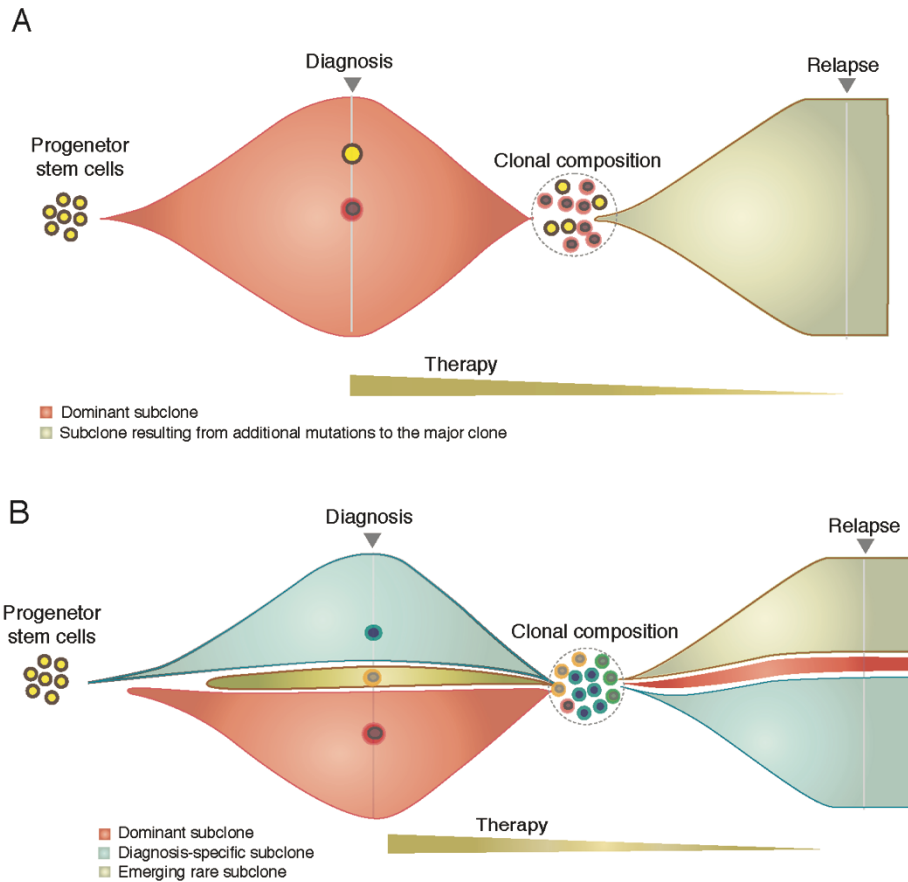


Figure 4: Clonal evolutionary trajectories of AML from early progenitor cells to later relapses. A) Linear evolution of cancer: A single subclone is dominant at diagnosis and relapse and, B) Branched evolutionary model: Multiple subclones arising from early progenitor cells undergo selective pressures induced by factors such as therapy. As a consequence of therapy, subclones dominant at diagnosis are lost at relapse (red) or rare subclones at diagnosis emerge as therapy resistant subclones dominating at relapse (yellow) or remain at the same frequency from diagnosis to relapse (blue). Illustration by Christian Benner. Adapted by permission from Springer Nature Publisher: Nature, “Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing” Li Ding et al.,⁴⁸ copyright 2012.

Temporal and spatial collection of tumor samples enables in-depth query into the dynamics of clonal evolution during disease progression. Studies on intra-tumoral heterogeneity of cancers between spatially distant primary and metastatic tumors have revealed branched evolution of the tumors, for example in renal cell carcinomas⁴⁴, prostate cancers⁷¹, breast carcinomas^{70,72} and lung cancers⁷³. Ding et al, comprehensively illustrated clonal evolution of AML patients from diagnosis to later relapse based on whole-genome sequencing of serially acquired tumor samples.⁴⁸ The study showed how therapy could select for a previously existing low-frequency subclone to dominate the cancer cell fraction at relapse. Preleukemic stem cells in leukemia have shown to contribute to both resistance and relapse of cancers.^{26,74,75} Similarly, low-frequency subclones pre-existing at diagnosis have been shown to re-emerge and seed metastasis in a distant tissue in prostate cancer.⁷¹ The emerging picture of tumor complexity due to subclonal evolution and heterogeneity can explain the failure of chemo and targeted drugs to truly eradicate cancer which so far has not taken subclonal heterogeneity and evolution into consideration.

Recent studies have proposed additional models of tumor evolution such as punctuated evolution, wherein large genomic aberrations accumulate in a short span of time either at chromosome or whole-genome level.⁷⁶⁻⁷⁹ This model of evolution has been referred to as the ‘big bang’ model leading to an evolutionary leap.⁸⁰ Phenomenon’s such as chromothripsis, kataegis, chromoplexy, which cause large-scale genomic instabilities, can result in such a model of evolution.^{77,79,81,82} These are known to arise in multiple cancer types, but are not often seen in leukemias.

2.3 Intra-tumor heterogeneity of cancers

Branched evolution underlying intra-tumor heterogeneity between primary and metastatic sites in renal cell carcinomas revealed the extent of ITH in cancers.⁴⁴ ITH in cancer is not a new discovery; morphological heterogeneity of cancer was recorded as early as 1800s.^{83,84} Although the concept of ITH evolved with growing understanding of cancer biology such as discovery of genes, drivers and histopathological advancement, not until the recent advances in high-throughput sequencing technologies in-depth enquiry into ITH has been possible.⁸⁵⁻⁸⁸

Sequencing-based interrogation of ITH revealed branched evolution of cancer subclones wherein both ubiquitous mutations and site-specific mutations were identified.^{44,46,51,72,89} Metastatic sites in solid tumors are genetically distinct from the primary tumor tissue with subclones emerging specifically in metastasis. Phylogenetic analysis of ITH has revealed that in majority of cancers trunk mutations with most recent common ancestors are shared among spatially distinct sites.^{44,70,90} The cancer evolves by further branching out and as a consequence of evolution, mixed cell populations are generated leading to ITH. Despite the extent of ITH in cancers, recent studies are providing evidence of convergent evolution in tumors wherein, spatially and temporally distinct subclonal mutations converge to the same gene or pathway eventually.^{78,91} Evidence of convergent evolution has been seen in many cancer types where multiple mutations are found in a single gene or in genes belonging to the same pathway.^{46,78,92} Targeting these pathways could provide a strategy for exploiting ITH for the benefit of treating cancer.

3. CLINICAL IMPACT OF TUMOR EVOLUTION AND HETEROGENEITY

Clonal evolution and heterogeneity has a profound impact on every aspect of clinical management of cancer such as diagnosis, treatment, disease management, trial designs, clinical follow up and drug discovery. Although, many recent studies have thrown light on the clonal complexity and heterogeneity of tumors, there has been minimal progress in constructing effective clinical trials and drug discovery processes to address the issue. Understanding the impact of therapy at a subclone level and interdependencies between multiple subclones will help us design strategies to tackle evolution and heterogeneity for the treatment of cancer.

3.1 Impact of therapy on clonal evolution

Chemotherapy is the first-line treatment strategy in hematological cancers.⁹³ Some of the chemotherapeutic agents act by disrupting cell division, inhibiting function of microtubules, interfere with DNA replication (antimetabolites) and inhibit topoisomerase enzymes. The indiscriminate and often genotoxic nature of chemotherapy results in a multitude of side effects including genomic instability, killing of the normal cell components including immune system, and could therefore contribute to the accelerated evolution of cancer.

Studies in relapsed AML showed how chemotherapy not just failed to eradicate the founding clone, but led to the selection of chemotherapy resistant subclones that dominated at relapse.⁴⁸ An increase in the number of transversions at genome-wide level was observed as a consequence of therapy as well. The ability of chemotherapy to induce random mutagenesis leads to heterogeneous cancer cell populations resistant to chemotherapy.⁹⁴ Similar effects of chemotherapy on clonal selection and expansion have been reported in myelomas,^{95,96} lymphomas,⁹⁷ gliomas⁹⁸⁻¹⁰⁰ and lung cancers.¹⁰¹ Targeted therapies have also been shown to select for therapy-resistant subclones.¹⁰²⁻¹⁰⁴ Targeting a single vulnerability has usually resulted in incomplete eradication of subclones and re-emergence of cancers resistant to the therapies used. Current targeted therapies in most solid and liquid cancers are usually targeting known recurrent “actionable” alterations without taking heterogeneity into consideration. Thus,

presences of tumor heterogeneity and cells intrinsically resistant to the drugs often leads to short-term remissions at best.

3.2 Emergence of drug resistant clones

Based on current knowledge, therapy can result in emergence of drug resistant subclones in the following ways: i) selection of a pre-existing drug-resistant subclone; ii) random mutagenesis leading to formation of novel drug resistant clone, iii) differential sensitivity of subclones leading to incomplete eradication of cancer cells or; iv) incomplete eradication of cancer stem cells. Numerous studies have observed the pattern of selection for resistant subclones by therapy.^{98-100,102-105} A classic example of this case is imatinib, a small molecule tyrosine kinase inhibitor of BCR-ABL achieved tremendous success in treating chronic myeloid leukemia (CML). Although initially responsive, many advanced-stage CML patients relapsed by acquiring resistant mutations in BCR-ABL rendering treatment with imatinib ineffective.^{103,104} Differential sensitivity of subclones to therapy resulting in emergence or outgrowth of new subclones have been reported by Landau et al and in our studies (paper I) Ojiamies et al.^{106,107}

3.2.1 Clonal and subclonal mutations

For the better understanding and targeting of drug resistant subclones, it is pertinent to classify clonal and subclonal mutations in the context of evolution. Mutations arising from early ancestral cell and present in all cancer cells are considered to be clonal or trunk mutations (Cancer cell fraction = 1). Whereas mutations present in a fraction of tumor cells are considered to be subclonal.⁶⁰ Targeting of trunk mutations alone, that is clonal mutations present in majority of the cancer cells, will result in emergence of low-frequency minor subclones at relapse or metastasis not targeted by the therapy. Also, for targeted therapy we need to consider if the targeted genes are clonal or subclonal. If the targeted genes form a minor fraction of cancer cells, this would result in ineffective therapy response. Landau et al. observed shorter event-free survival rate after therapy in patients with subclonal mutations when compared to patients without subclonal mutations. Therefore, subclonal mutations are an independent risk factor of relapse and disease progression.¹⁰⁶ Given that pre-existing low-frequency subclones are seen to emerge at relapse, it is important to gain a comprehensive

understanding of the clonal composition and design therapies targeting multiple subclones.

4. LONGITUDINAL AND SPATIAL SAMPLING OF TUMORS

The dynamic nature of cancer evolution, therapy resistance and intra-tumor heterogeneity emphasizes the importance of sampling tumors during the course of the disease from multiple sites. Tumor sampling from a single site underestimates the extent of intra-tumor heterogeneity in tumors. Samples are typically taken in the clinic at diagnosis and infrequently during the course of the disease. Single time point sampling provides only a snapshot of alterations that have continually undergone evolutionary selection. Further, impact of therapy on the clonal heterogeneity and evolution of cancer is not captured from single time point samples limiting our understanding of drug resistance and therapy response.

Therefore, an enhanced sampling strategy to include both longitudinal and spatially acquired samples from primary and metastatic tumors needs to be adopted. Although this approach is not clinically feasible, adoption of strategies such as sampling circulating tumor cells or DNA multiple times before, after and during therapy can shed light on emerging subclones, resistant mechanisms and mitigate the impact of ITH. Many recent studies have shown the validity of this approach in studying ITH and therapy response.¹⁰⁸⁻¹¹¹

AIMS OF THE STUDY

The overall aim of the thesis was to comprehensively analyze clonal evolution and heterogeneity of cancer subclones and to delineate the impact of therapy at subclone-level in the context of individualized cancer treatment. For these studies, acute myeloid leukemia (AML) and renal cell carcinoma (RCC) –derived *in vivo* patient samples and patient cells in *ex vivo* culture, were analyzed. .

The specific aims were:

1. Develop improved methods to determine the clonal evolution of cancer during chemo- and targeted therapies, with a focus on low-frequency subclones in patients receiving treatment for AML.
2. Investigate genetic heterogeneity between different bone marrow compartments in AML.
3. Apply patient-derived cell cultures (PDCs) to investigate the evolutionary patterns in renal cell cancer, including subclone-specific drug responses.

MATERIALS AND METHODS

Material and methods used in this thesis are described in detail in the original publications (I-III). Below is the summary of materials and methods in brief.

1. PATIENT STUDY MATERIAL – AML (I, II)

For studies I and II, bone marrow aspirates (BM), peripheral blood (PB) and skin samples (germline control for sequencing experiments) were obtained from AML patients. Prior informed patient consent (Permit numbers 239/13/03/00/2010, 303/13/03/01/2011) in accordance with the Declaration of Helsinki was received from Helsinki University Hospital Ethics Committee.

For Study I, a total of 39 longitudinal BM samples from 12 AML and one high-risk myelodysplastic syndrome (MDS) patient were collected. For Study II, mononuclear cell fractions of BM samples of two AML patients were obtained from three distinct sites each: corpus sternum (CS), right and left iliac crest (RIC and LIC). Blast counts from bone marrow aspirates are detailed in respective publications. Mononuclear cell fraction was isolated from patient bone marrow and blood samples using density based gradient separation (Ficoll-Paque Premium; GE Healthcare, Little Chalfont, UK). Genomic DNA from the mononuclear cells and skin sample were isolated using DNeasy Blood and Tissue kit (Qiagen, Hilden, Germany).

2. STUDY MATERIAL – RCC (III)

2.1 Patient samples

Tumor and benign tissue samples of four renal cell cancer patients were obtained from primary and metastatic sites. Histopathological characterization of benign and malignant cells was performed using hematoxylin and eosin staining. Signed informed consent (Dnro 263/13/03/02/2011; 379/13/03/02/2012 and Dnro § 212, approved by Urological Biobank Initiative with Helsinki Urological Biobank; HUB) was obtained from patients in accordance with the Declaration of Helsinki.

2.2 Patient derived cultures (PDCs)

Patient derived cultures from primary cells were established from conditionally reprogrammed cells (CRC). Hematoxylin and eosin staining was used to evaluate the ratio of benign and malignant cells from parental tissues to characterize tumor and benign tissue cells. Detailed description on method to establish the PDCs are described in the original publication III.

3. GENOME SEQUENCING TECHNOLOGIES

3.1 Whole exome sequencing analysis (I – III)

Genomic DNA extracted from tumor, skin (I, II) and benign tissue (III) cells were used for whole-exome sequencing (WES) analysis. Initial exome capture on 3µg of DNA was performed using Agilent SureSelect Human All Exon V5 (Agilent, Santa Clara, CA, USA) or Nimblegen SeqCap EZ v2 capture kit (Roche NimbleGen, Madison, WI, USA) and sequencing was performed using Illumina Genome Analyzer Iix, HiSeq1500 or 2500 instruments (Illumina Inc., San Diego, CA, USA). On average we achieved 30x coverage for normal control and 100x coverage for tumor samples. Somatic mutations, i.e. tumor-specific variants and variant allele frequencies (VAF) were determined using VarScan2 algorithm¹¹² with the following parameters:

- Strand-filter 1
- Min-coverage-normal 8
- Min-coverage-tumor 1
- Somatic-P value 0.01
- Normal-purity 0.95
- Min-var-freq 0

Mutations were annotated with SnpEFF¹¹³ using Ensemble v68 version. Cancer cell fraction was calculated to estimate tumor purity (in I) as described by Rubio-Perez et al.¹¹⁴ The VAF were corrected for the cancer cell fraction. By doing so we were able to correct for sample variation in blast counts and copy number variations.

Copy number variations (CNV) were determined using in-house CNV analysis pipeline. Briefly, we calculated RPKM (reads per kilobase of target region length per million mapped reads) value for each target region from both tumor and skin control samples. Target regions with <25x coverage were excluded from analysis and relative log₂ ratio of copy number calls for the tumor versus skin were calculated and segmented using Circular Binary Segmentation.¹¹⁵ Copy number changes were set at -0.4 (heterozygous deletion), -1.2 (homozygous deletion), +0.5 (gain) and +1.3 (amplification).

3.2 Ultra-deep amplicon sequencing (I – II)

Putative, clonal and subclonal driver mutations identified from WES data were validated using ultra-deep amplicon resequencing. Also, variants that were identified specifically at certain time points (diagnosis, relapse) were validated in all serial patient samples to determine if the variants were present at low-frequency (Figure 5A).

We performed amplicon resequencing using Illumina MiSeq instrument with MiSeq Control Software v2.4 or newer (Illumina, Inc., San Diego, CA, USA). Initially target sequence primers were prepared using in-house targeted PCR amplification protocols. The protocol used either one or two rounds of PCR amplification. In the two-step protocol, first amplification round was performed using 10-20 ng of sample DNA, 10 µl of 2x Phusion High-Fidelity PCR Master Mix (Thermo Scientific Inc., Waltham, MA, USA), 0,5 µM of each locus specific primer carrying Illumina-compatible adapter sequences at annealing temperature 60 °C. Subsequently, samples were purified using Performa® V3 96-Well Short Plate (EdgeBio, Gaithersburg, MD, USA) and QuickStep™2 SOPET™ Resin (EdgeBio, Gaithersburg, MD, USA) according to the manufacturer's instructions. A second round of PCR amplification was performed to include index primers carrying Illumina grafting P5/P7 sequence.

Subsequent to PCR amplification, additional purification steps were performed with Agencourt® AMPure® XP (Beckman Coulter, Brea, CA, USA) according to the manufacturer's instructions. We quantified amplification performance and amplicon yield on Agilent 2100 Bioanalyzer using Agilent High Sensitivity DNA Kit (Agilent Technologies Inc., Santa Clara, CA, USA). Samples were sequenced

as 251 bp paired-end reads and two 8 bp index reads on MiSeq sequencing platform. Further analysis on sequenced data was performed using an in-house amplicon analysis pipeline.

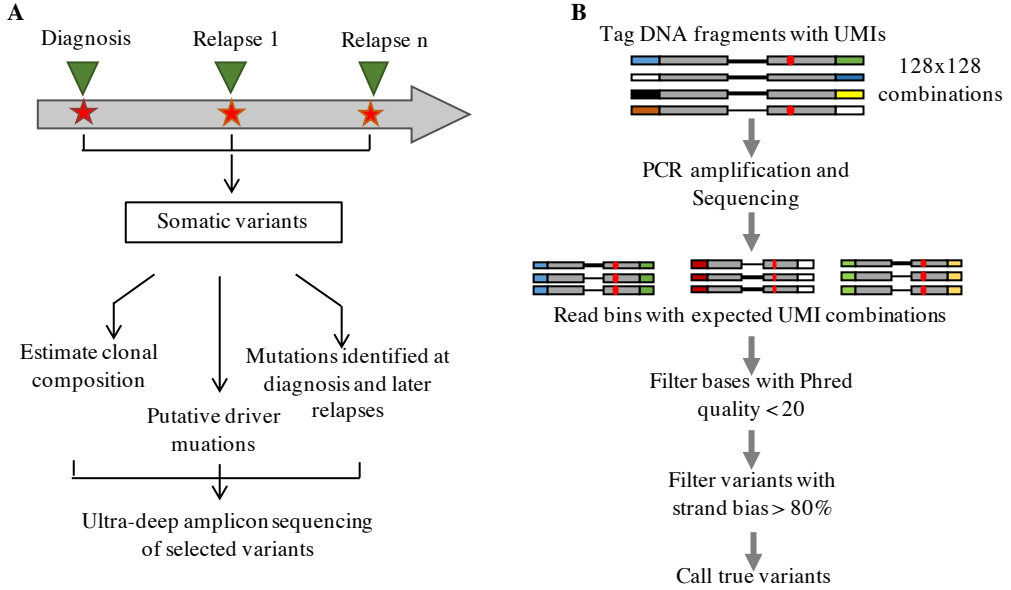


Figure 5: UMI-based ultra-deep amplicon sequencing approach. A) We estimated clonal composition, driver mutations and diagnosis or relapse-specific mutations from serial samples acquired during the disease course. Mutations belonging to either of the three criteria were selected for further validation using ultra-deep amplicon resequencing. B) To enhance rare variant calling, we generated 16384 UMIs and tagged DNA fragments. Subsequent to PCR amplification and sequencing, reads with expected UMI combinations were selected for and further base quality filtering was performed. Variants were called from the resulting filtered data.

3.2.1 Unique molecular identifier (UMI) based ultra-deep amplicon sequencing (I, II)

We utilized UMI-based ultra-deep sequencing approach to enhance variant calling by reducing mismatch errors introduced by PCR amplifications, sequencing errors, low quality sequences, among others. We adopted a non-random approach for generating UMIs wherein 128 degenerate 6 bp sequences with an edit distance of 2 bp and non-complementary to the index sequences are

generated. This approach has two advantages. Firstly, by utilizing combination of non-random UMI's (128x128 degenerate sequences), we were able to have control over sequencing errors in the barcode regions and have prior knowledge of the expected UMI combinations. Secondly, in our case with low sample amounts, UMI's enabled efficient multiplexing and increased sensitivity of low variant calls in a cost-effective manner. In order to remove contaminating germline variants, we performed amplicon sequencing on matched skin control and normal DNA control samples. Rigorous data filtering steps were implemented to reduce mismatch and PCR introduced errors and enhance true variant calling (Figure 5B).

4. DRUG SENSITIVITY AND RESISTANCE TESTING (DSRT) (I, III)

We performed *ex vivo* DSRT on patient-derived bone marrow mononuclear cells (in Study I) and patient-derived cell cultures (PDC's in Study II) with 306 clinically approved and emerging oncology drugs as described by Pemovska et al¹¹⁶. Cells were incubated in five different drug concentrations for 72 h. Subsequently cell viability was determined with CellTiter-Glo (CTG, Promega, Madison, WI, USA). Dose response curves were generated for each drug using Dotmatics Studies software (Dotmatics Ltd., Herts, UK). We quantified drug response of each drug in a sample by estimating drug sensitivity score (DSS). The DSS score was estimated as described by Yadav et al.¹¹⁷

5. STATISTICAL ANALYSIS (I-III)

Statistical analyses were performed with R statistical software (<http://www.r-project.org/>) and Prism software version 6.0 (GraphPad Software, San Diego, CA, USA). Correlation analysis between two variables was assessed using Pearson's correlation (two-tailed) or Spearman's rank correlation analysis. Linear regression method was applied to fit longitudinal data points collected during disease progression.

5.1 Correlation matrices (I, II)

For spatially and temporally acquired tumor samples, we developed a correlation-based method to visualize the subclonal composition of the patient. Briefly, based on correlation scores of VAF's across multiple time points, we quantitatively modelled clonal composition. Using trend estimation based on linear regression models, we tracked subclonal response to therapy. Based on these two estimates, we were able to delineate mutations from longitudinal samples into specific subclones. Significance of subclonal response to therapy was determined based on χ^2 test of absolute variant reads of sample taken before and after treatment. A positive correlation between mutations indicates that they are part of the same subclone, whereas a low or negative correlation score indicates the probability of the two mutations forming part of distinct subclones.

6. PHYLOGENETIC ANALYSIS (II, III)

We reconstructed cancer phylogeny to understand the clonal relation and progression among multiple tumor samples using the method published by Popic et al.¹¹⁸ Briefly, the method robustly groups the samples based on the presence or absence of mutations in a subgroups. These groups are further clustered using expectation-maximization clustering algorithm for Gaussian mixture models. Phylogenetic trees are constructed based on maximum parsimony method. We implemented the tool to study cancer evolution from anatomically distinct bone marrow compartments in Study II and between primary, metastatic tumors and respective PDC's in Study III. The following parameters were used for the analysis:

- maxVAFAbsent 0
- minVAFPresent 0.0001
- minClusterSize 4
- minPrivateClusterSize 2
- maxVAFValid 0.8

RESULTS

1. MONITORING THERAPY RESPONSE AT SUBCLONE-LEVEL

1.1 Personalized medicine platform for analyzing subclonal therapy response

As part of our individualized systems medicine (ISM) program, therapies are tailored for chemorefractory AML patients based on integrated information from genomic profiling and functional drug response and resistance profiling (Figure 6).

In Study I, we analyzed longitudinal samples taken at diagnosis, relapses, before and after therapy from 13 chemorefractory AML patients. Six of the 13 patients received targeted therapies based on the results from the ISM approach. This enabled us to monitor subclonal responses to both chemotherapy (13 patients) and to targeted therapy (six patients). Initially, we performed WES to identify somatic variants and determine clonal composition of each patient sample. Subsequently we performed ultra-deep amplicon resequencing with UMIs to validate subclonal mutations as well as to identify rare variants not detected by exome sequencing approach.

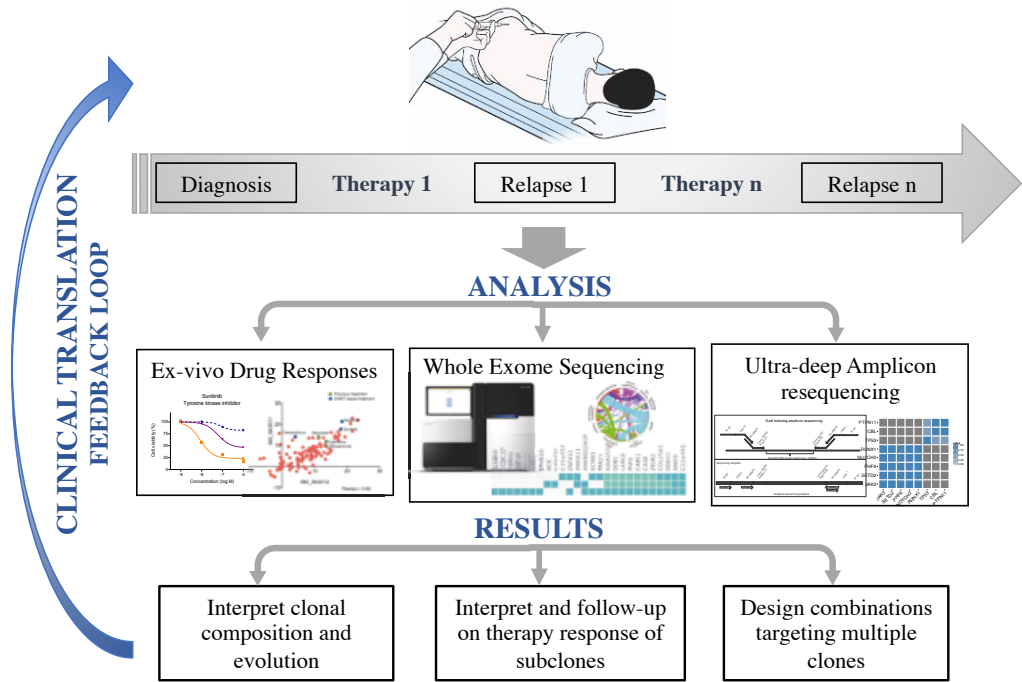


Figure 6: Overview of clonal evolution and response to therapy approach as part of the individualized systems medicine (ISM) program. Serial tumor samples obtained at diagnosis and during disease progression were subjected to DSRT, exome-sequencing and ultra-deep amplicon sequencing analysis. Based on the obtained data, we interpreted clonal evolution, heterogeneity and therapy response of individual subclones that enabled us to design combinatorial therapy regimens to target multiple subclones. The information gained is translated to the clinic, thereby providing an effective feedback loop between research and clinical translation.

1.2 Methodologies for interpreting subclonal architecture and therapy response from multiple time point samples

In this study, we developed improved methodologies to: i) enhance the detection of rare subclonal variants (frequency < 10%), ii) comprehensively estimate subclonal composition from serially acquired samples and iii) assess the impact of therapy on individual subclones over multiple time points.

To enhance detection of low frequency variants, usually undetected by whole exome or genome sequencing efforts due to low signal-to-noise ratio we developed an ultra-deep amplicon resequencing with UMI based approach. The

method enabled sensitive detection of low-frequency variants by achieving very high coverage ($>10000\times$). By utilizing UMIs in addition to relevant filtering steps we were able to significantly minimize the noise, while enhancing the sensitivity to detect low-frequency variants. Using this approach, 16 low frequency variants (0.5% - 2%) could be identified, that were undetected by WES. In one such case, we identified *PHF6* variant in patient #706 present at low frequency. The variant was seen at 20% frequency at relapse but was not detected at diagnosis. We could validate the variant to be present at 0.5% frequency at diagnosis, by significantly reducing the background noise at such low frequency levels (Figure 7).

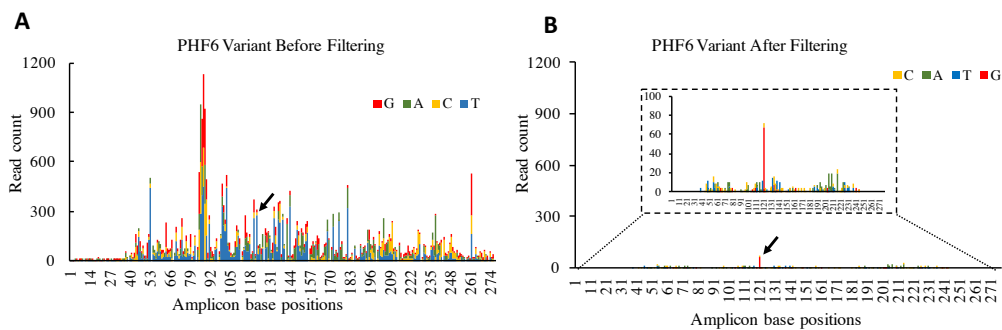


Figure 7: Before (A) and after (B) implementation of UMI-based filtering approach to optimize variant calling of low-frequency *PHF6* variant. A significant reduction in the background noise level is shown after implementing the filtering while validating the authenticity of the *PHF6* variant to be true. Inset in B shows the *PHF6* variant in detail.

Secondly, to estimate subclonal composition from temporally acquired tumor samples, we implemented a correlation-based method wherein we built correlation matrices (Pearson correlation) for each variant based on their allele frequencies in different time points. The likelihood of each variant belonging to a subclone was determined based on the correlation coefficients. A positive correlation between mutations indicates that they are likely to form part of the same subclone, whereas a low or negative correlation score indicates the probability of the mutations forming part of distinct subclones. Using this approach, we quantitatively modelled clonal composition. For example, in patient #1145 we identified here distinct subclones based on their correlation scores (Figure 8A). Mutations in *JAK3*, *STED2*, *PHF6*, *NOTCH2* and *RUNX1* had very high correlation scores compared to each other indicating that these

belong to the same subclones. However, these subclonal mutations had a negative correlation score with *TP53*, *CBL* and *PTPN11* indicating that these mutations belong to a separate subclone (Figure 8A).

To estimate the impact of subclones on therapy response, we performed trend analysis using linear regression models. With the availability of longitudinal samples before and after therapy, linear trend estimation helped us identify if mutations arising at later time points belonged to the same subclone or a different one. For example, in patient #1145, by comparing variant allele frequency of *TP53* against other mutations, we saw a linear trend of subclonal variants segregate. Subclone with *PTPN11* and *CBL* showed a positive slope whereas the other subclonal variants had a negative slope in correlation to *TP53* (Figure 8B). Additionally, to determine the significance of subclonal response to therapy, we performed χ^2 test of absolute variant reads of sample taken before and after treatment.

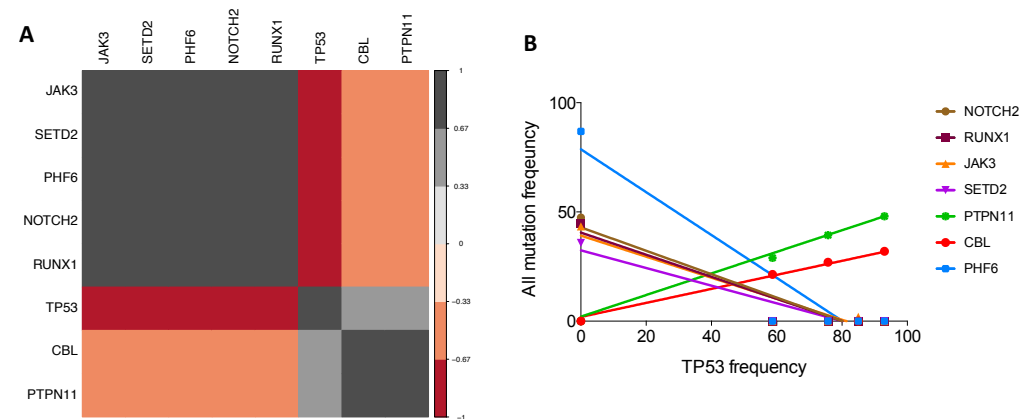


Figure 8: A) Correlation matrix indicating subclonal composition of the leukemic cells in patient #1145. Genes with grey block highly correlated among each other whereas red blocks indicate negative correlations. We saw at least three distinct subclones based on these mutations. B) Linear trend analysis indicated the trend of all subclonal mutations in comparison to *TP53* mutations. In this case, we saw that the subclones were each responding differentially in comparison to each other.

1.3 Subclonal responses to chemotherapy

We observed responses of individual subclones to chemotherapy in 13 AML patients who were treated with a combination of cytarabine and idarubicin. We

compared untreated diagnostic samples against samples taken after therapy to study i) differential response and resistance of subclones to chemotherapy and ii) emergence and loss of subclonal mutations subsequent to therapy. Based on chi-square trend test we found statistically significant response ($p < 0.0001$) of individual subclones to chemotherapy. Interestingly, we discerned either emergence or loss of subclonal mutations in response to therapy. In patient #600 we observed emergence of new mutations as well as selection of existing mutations in *WT1* gene after chemotherapy (Figure 9A). The data obtained implies the selective pressure of chemotherapy on existing subclones and the rise of new, potentially cancer driving, mutations. In patient #706, we observed three different subclones responding in a completely different fashion (Figure 9B). A dominant subclone with driver mutations in *DNMT3A* and *TET2_1* responded only marginally to therapy. Whereas, subclone with *NPM1* and *TET2_2* was almost completely eradicated after chemotherapy. Thirdly, we saw emergence of a novel subclone with *PHF6* mutation at relapse, which was present as a minor subclone at diagnosis (0.5%). This case illustrates the differential effect therapy could have on subclones and how chemotherapy acts as a bottleneck leading to subsequent evolution of subclones.

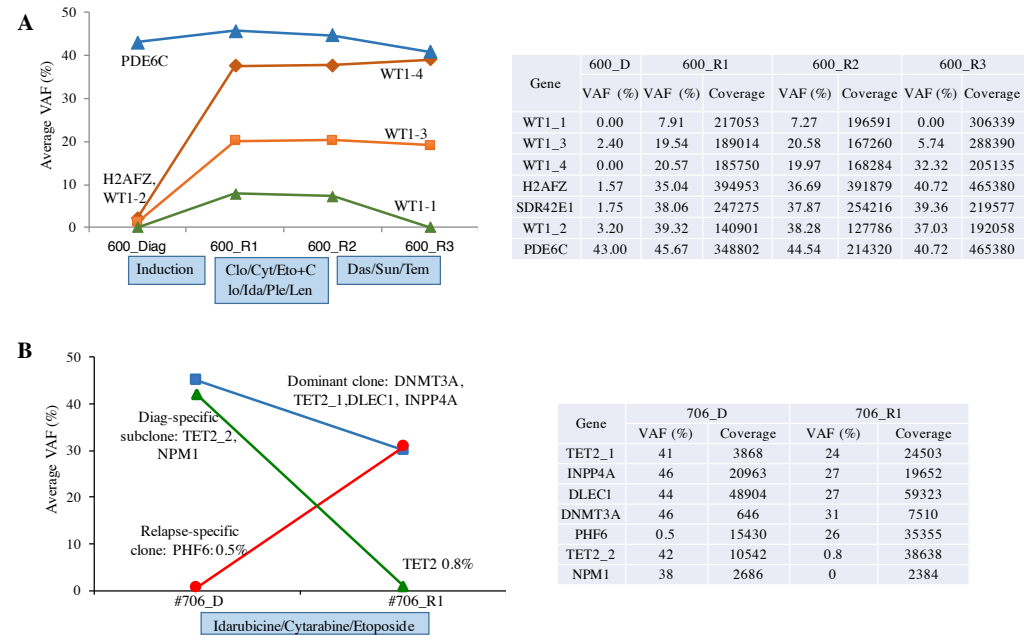


Figure 9: A) Subclonal dynamics of the leukemia cell population in patient #600 from diagnosis to relapse. We see new emerging subclonal mutations in *WT1* at first relapse (600_R1). *WT1*-1 mutation responded to therapy and was eventually eradicated but *WT1*-3 and *WT1*-4 were resistant to therapy indicating that they belong to a different subclone and have acquired resistance to therapy. B) In patient #706 we see differential response of subclones to chemotherapy; 1) a dominant clone that remained unaffected by therapy, 2) a diagnosis – specific subclone that responded to therapy and was not seen at relapse, and 3) a subclone with *PHF6* mutations reemerging at relapse as a dominant clone from a pre-existing low-frequency subclone at diagnosis.

1.4 Subclonal responses to targeted therapy

Six of the 13 patients were given individually tailored targeted therapies based on the ISM approach. We aimed to delineate the response of targeted therapy on individual subclones from samples taken before and after therapy. Based on clonal variant frequencies in five of the six patients, we observed significant reduction in clonal frequencies indicating response to targeted therapy ($p < 0.0001$).

Patient #560 received temsirolimus-dasatinib targeted therapy based on drug response profiling data. Objective clinical response based on the European LeukemiaNet (ELN) criteria¹¹⁹ was not observed for this patient. We acquired five serial samples from diagnosis to later relapse, including samples taken during, before and after the targeted therapy (Figure 10A). We identified a dominant subclone with *GATA2DB* and *CBFB* mutations at diagnosis (#560_D), which then acquired additional mutations of *BRCA1*, *PTPN2*, *PTPN11* and *MLH1* genes, detectable at relapse (#560_R1). An additional subclone with *CBL* mutation was seen at relapse (Figure 10B). Subsequent to given targeted temsirolimus-dasatinib therapy we observed a significant quantitative response to therapy for both subclones during therapy, with the average VAF decreasing to 4% ($p: 0.003$). The patient however relapsed quickly, reflected also by the increasing subclonal frequency (average VAF: 21%) of the sample taken at relapse (#560_R2). This indicates that the two subclones acquired resistance to therapy (Figure 10C). Although the patient showed no response by objective clinical evaluation based on ELN criteria, a transient quantitative response at the subclone level could be observed during the therapy.

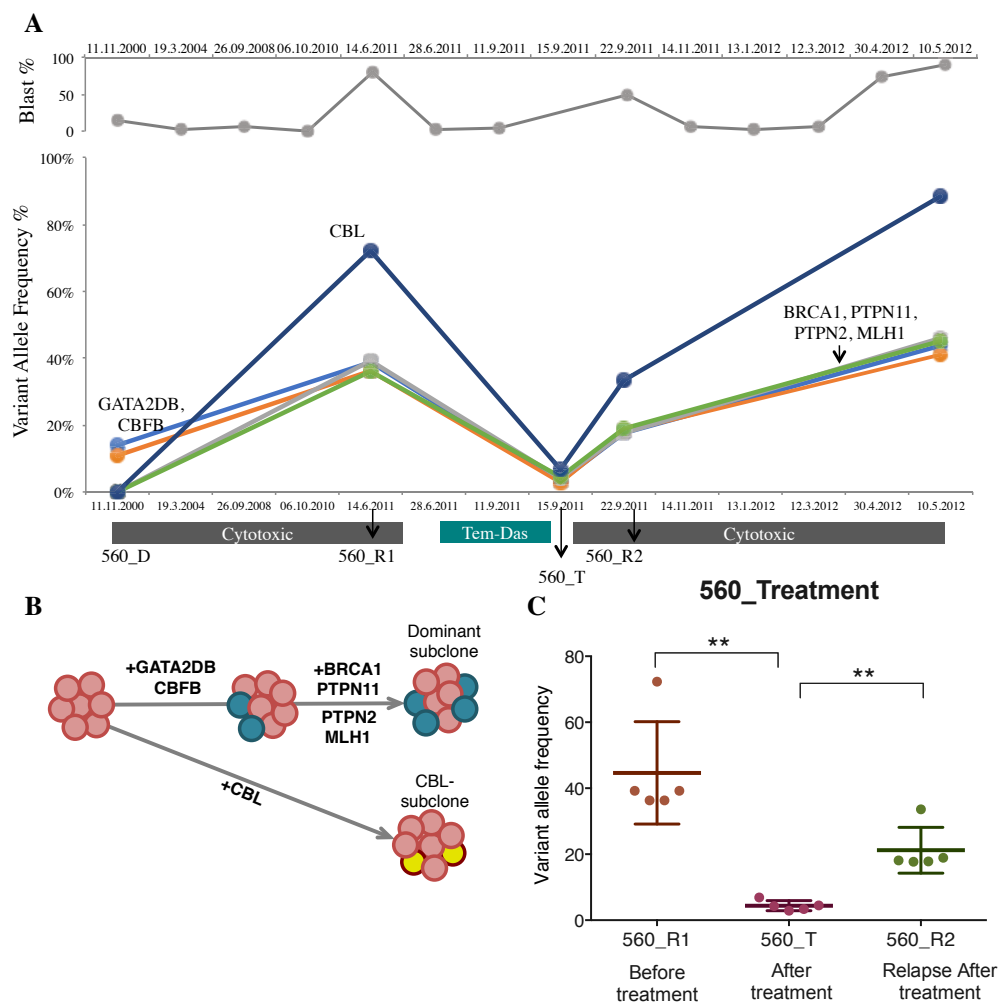


Figure 10: (A) The disease course and frequency of subclones in patient #560 at five different time points. Blast percentages are shown on top. Patient received chemotherapy and targeted temsirolimus-dasatinib -based therapy. In this patient, two subclones with mutations affecting *CBL* in one clone and *BRCA1*, *PTPN11*, *PTPN2* and *MLH1* in other, were gained at first relapse. We observed both subclones responding to therapy as indicated by decreased subclonal frequency at 560_T and subsequently the subclones acquired resistance to therapy and dominated at relapse (560_R2). (B) Schematic representation of evolution of AML in #560. Based on the VAF over multiple time points we can hypothesize that mutations affecting *BRCA1*, *PTPN11*, *PTPN2* and *MLH1* were acquired in the same cells with *GATA2DB* and *CBFB* mutations, resulting in a dominant subclone detected at relapse (560_R1). VAF of *CBL* indicated that it constituted a

separate subclone and that at 560_R1 majority of the cells were seen carrying *CBL* mutation. (C) Plot depicts the VAF of subclonal variants in samples taken before therapy (560_R1), during therapy (560_T) and after therapy (560_R2). We see a significant decrease subsequent to therapy and an increase in frequency after therapy indicating emerging resistance (χ^2 test, $P = .003$).

For patient case 1886, serial samples from diagnosis and three subsequent relapses were analysed after chemotherapy and individually chosen targeted therapies (Figure 11A). Clonal analysis revealed a major clone with *DNMT3A*, *FLT3* and *NPM1* mutations. As illustrated, mutations in *DNMT3A* were followed by acquisition of *FLT3* mutations (Figure 11B). Based on the VAFs we could determine *DNMT3A-FLT3* mutated subclonal cells as well as non-mutated cells acquiring *NPM1* mutation, resulting in two distinct subclones. Additionally, a minor subclone with *STAT5A*, *MTOR* and *CBL* mutations, not detected at diagnoses (despite of an average sequencing coverage >11000x), emerged at relapse. Since the patient received a bone marrow transplant, we cannot ascertain if the minor subclonal mutations were cancer-specific or donor-derived polymorphisms. *KIAA1209-PDGFRB* fusion gene was detected at diagnosis and relapse, but could not be followed up due to unavailability of RNA sample.

Subsequent to first relapse (#1886_R1), the patient received azacitidine and sunitinib therapy based on the presence of *PDGFRB* fusion gene as well as *FLT3* mutations which both link to sunitinib sensitivity.¹²⁰⁻¹²² Furthermore, ex-vivo sensitivity verified the responsiveness to sunitinib. Subsequent to azacitidine/sunitinib treatment, the patient showed a complete clinical response and we observed a significant response at subclone level ($p < 0.0001$). Following this, the patient received targeted treatments with dasatinib and axitinib-everolimus but did not exhibit a clinical response. Although, we observed significant response of the minor subclone with mutations affecting *MTOR-CBL-STAT5A* to therapy, similar response was not exhibited by the other subclones (Figure 11A).

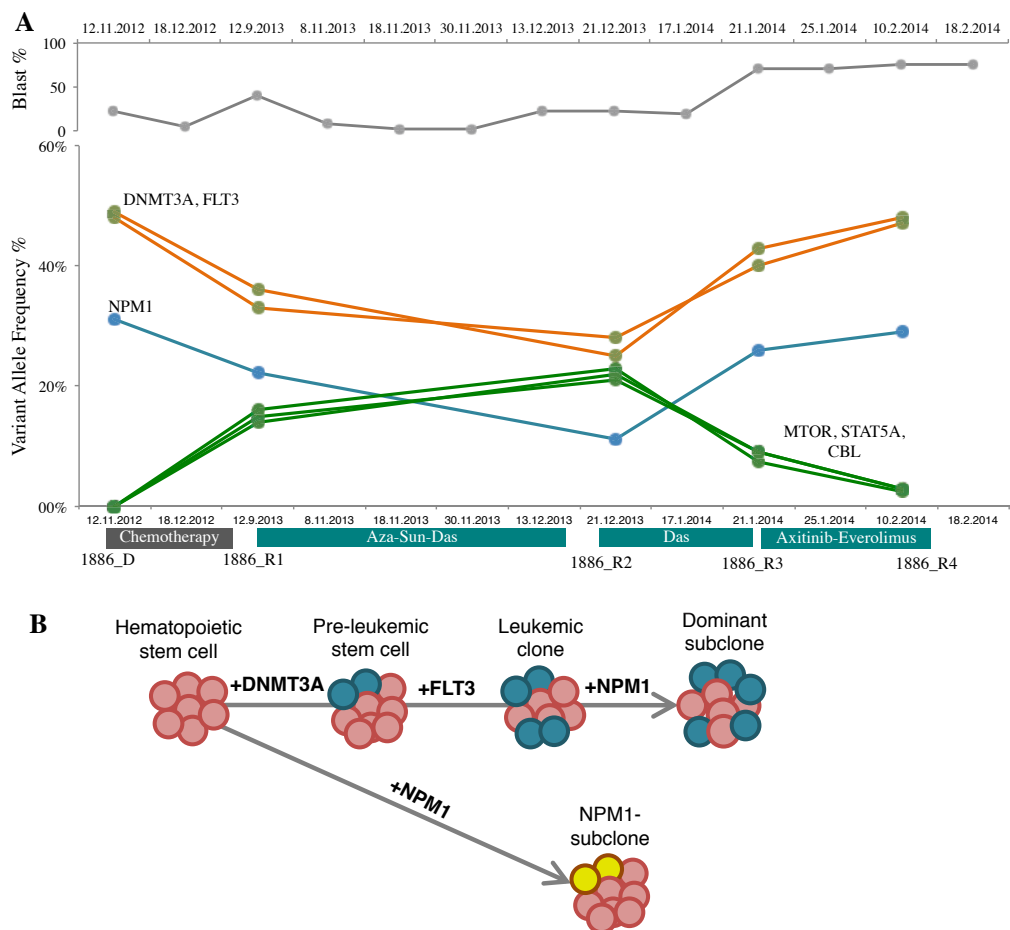
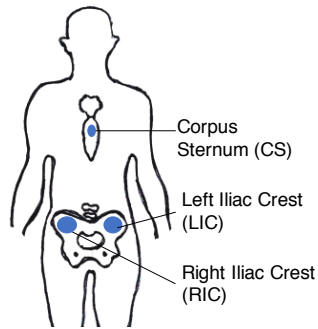


Figure 11: (A) The disease course and frequency of subclones in patient #1886 at five different time points. Patient initially received chemotherapy and subsequently individually tailored therapy based on *ex vivo* drug response profiles. Shown in the plot are responses of the different subclones to therapy over multiple time points. We observed that both chemo and targeted therapies did not have high impact on the *DNMT3A*, *FLT3* and *NPM1* – mutated subclone, resulting only in transient responses to targeted therapies tailored for the patient. (B) Schematic representation of evolution of AML in #1886. Mutation in *DNMT3A* represents an early pre-leukemic event, followed by mutations in *FLT3* and *NPM1* in the same clone. Due to the difference in VAF of mutated *DNMT3A-FLT3* and mutated *NPM1* over multiple time points, it is likely that a new subclone with only mutated *NPM1* emerged in parallel.

2. GENETIC HETEROGENEITY BETWEEN SPATIALLY DISTINCT BONE MARROW SITES IN AML

In Study II, we investigated if genetic heterogeneity exists between anatomically distant bone marrow compartments of AML patients. For this study, we acquired bone marrow aspirates from three bone marrow sites: corpus sternum (CS), right and left iliac crest (RIC and LIC). We performed WES followed by validation with ultra-deep amplicon resequencing to study the underlying heterogeneity of each sample. For this study, we recruited two AML patients. Patient case AML1, aged 74, was first diagnosed with MDS and later relapsed with secondary AML. The patient received azacitidine-based treatment, during which we obtained the BM samples. Patient case AML2, aged 72, received cytarabine-idarubicine - based induction therapy subsequent to AML diagnosis. The patient relapsed after two years, at which point the BM samples were biopsied.

In patient AML1, exome sequencing data of multiple bone marrow samples was used to identify 42 high confidence somatic mutations (somatic p value < 0.01). Of these, 18 mutations were shared between RIC and CS, but were not detected in LIC. Subclonal composition of the leukemic blasts, inferred based on correlation coefficient matrices, identified three distinct subclones in this patient. Further, we inferred cancer cell lineage of the multiple BM samples from the common ancestor using phylogenetic analysis tool LICHeE.¹¹⁸ The phylogenetic analysis revealed the genetic relation between the three BM compartments and underlying clonal heterogeneity. All three compartments shared a common ancestral clone with 10 mutations in genes such as *C22ORF43*, *TRPV3* and *ANO9*. Subsequently, the BM sites evolved in parallel, wherein two different subclones arose from the ancestral cell (Figure 12). One of the subclones with seven mutations was shared among the three sites. The other subclone with 16 mutations, including variants in known driver genes such as *PHF6*, *CABS1* and *ZRSR2*, was acquired by cancer cells in RIC and CS only, resulting in the branching out of the two sites from LIC making them genetically distinct. Interestingly, the subclonal frequency varied significantly between the two sites (9% and 22% in RIC and CS respectively).



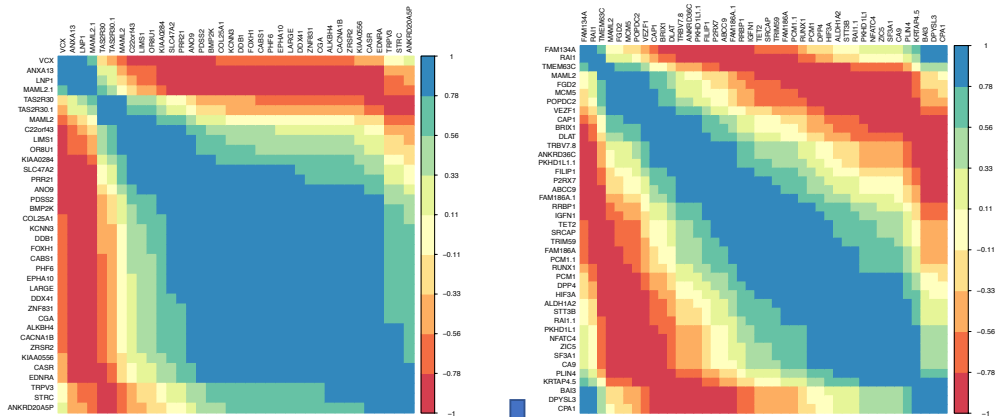
Exome sequencing



Estimate subclonal composition

AML1

AML2



Infer Cancer cell lineage

AML1

AML2

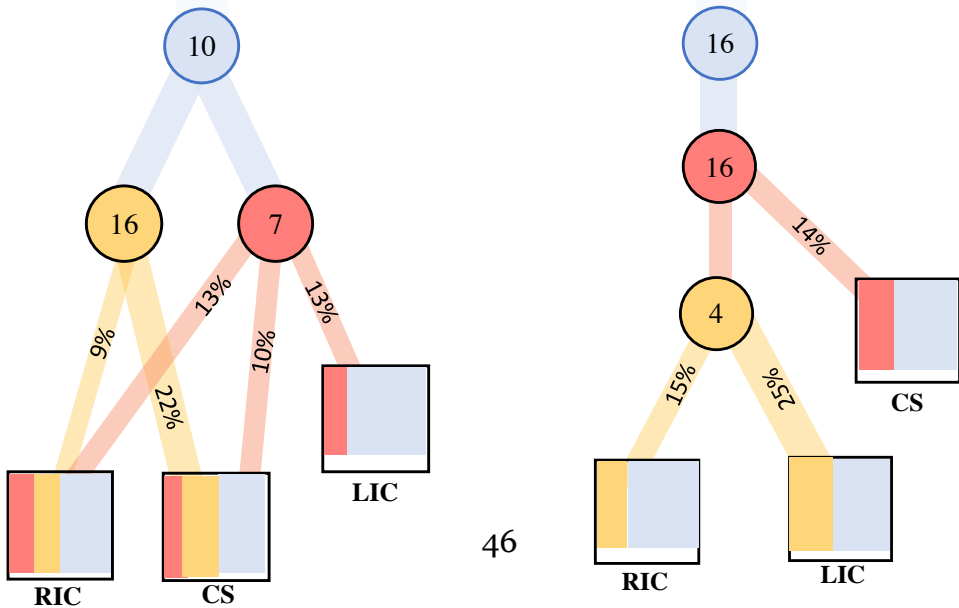


Figure 12: Inferring genetic heterogeneity between multiple bone marrow compartments of patients AML1 and AML2. We estimated clonal composition of the two patients using correlation-based analysis. As shown in the correlations matrices, we identified three major subclones in AML1 and at least four major subclones in AML2. Subsequent phylogenetic analysis of the two patients revealed genetic interdependencies of the multiple bone marrow compartments. Numbers indicated in the circles refer to the number of subclonal variants whereas subclonal frequencies in percentages are indicated on the branches. In AML1 we observed LIC to be genetically distinct from RIC and CS. The subclonal frequencies were significantly different between RIC and CS. Similarly, in AML2 we observed CS to be genetically distinct from RIC and LIC. Also, subclonal frequencies varied significantly between RIC and LIC.

In patient AML2, we identified 50 somatic mutations of which 40 were ubiquitous among all three sites, five mutations were shared between RIC and LIC and three mutations were shared between LIC and CS. We identified four distinct subclones in AML2 from correlation analysis (Figure 12). Cancer cell lineage inferred from phylogenetic analysis showed the first common ancestral clone with 16 mutations, including those affecting known driver genes *RUNX1*, *SF3A1* and *FILIP1*. The ancestral clone eventually acquired additional 16 mutations in all three sites. Subsequent to this, RIC and LIC specifically acquired four additional mutations including one in *SF3A1*, a known AML driver gene, resulting in genetic diversification from the CS site. Similar to AML1, we observed significant difference in subclonal frequency between the two sites (15% and 25% in RIC and LIC respectively).

3. GENETIC HETEROGENEITY UNDERLYING DRUG RESPONSES IN RENAL CELL CARCINOMA

In Study III we developed patient-derived cell cultures (PDCs) from benign, primary tumor and metastatic tumor tissues of four renal cell carcinoma patients. We performed exome sequencing of both the tissue of origin and PDCs for all patients and drug sensitivity and resistance testing for PDCs with 460 approved and investigational oncology drugs. The aim of the study was to i) infer clonal heterogeneity and evolution of primary, metastatic tumors and their respective PDCs, ii) analyze how representative the PDCs are of their respective tumor

tissues from which they were derived, and iii) study the impact of genetic heterogeneity on drug response profiles.

3.1 Genetic heterogeneity and relation between multiple tumor sites and established PDCs

Somatic mutations inferred from WES data were compared between tumor tissues and their respective PDCs of renal cell carcinoma (RCC) patients. Of the four cases (RCC.1-4), we found three patient models to be highly representative of their parental tissues (RCC.1-3) based on shared somatic mutations and copy number alterations (CNAs). In RCC.1 we identified alterations in known driver genes such as *VHL*, *PBRM1*, *KDM5C*, *TSC2* and *PIK3C2* in the tumor tissue, which were recapitulated also in the PDCs. Similarly, RCC.2 and RCC.3 parental tissues and PDCs shared common recurrent alterations such as deletions of chromosome 14 (affecting genes xx), 3p (*VHL*, *PBRM1*, *SETD2*, *BAP1*) and 9p (*CDKN2A/B*) and amplification of 7p (*EGFR*). RCC.4 PDCs did not share mutations with the parental tissue and therefore was not representative of the tumor.

To extend our understanding of the clonal relation between parental tissues and PDCs we performed phylogenetic reconstruction analysis to infer the cancer cell lineage among the different samples using LICHeE tool. In patient case RCC.1, consisting of four samples (tumor tissue, metastatic tissue and their respective PDCs), we observed a common ancestral cell clone with 36 mutations (Figure 13A). Subsequently, four additional mutations were acquired by the tumor PDC, metastatic tissue and metastatic PDC. These subclonal mutations were not identified in the tumor tissue. The metastatic tissue acquired 10 additional mutations resulting in a subclone specific to metastatic tissue, which was not seen in any other sample.

Similarly, for patient case RCC.3, with primary tumor tissue sample, two metastatic samples (vena cava and secondary metastasis) and their respective PDCs (six samples in total), we could reconstruct the phylogenetic tree based on the somatic mutations. In this case, a common ancestral clone with nine mutations was shared among all samples and was the only identified cancer clone in tumor

PDC. PDC derived from the vena cava site did not share any mutations and branched away from other samples. Subsequently, the ancestral clone acquired mutations sequentially resulting in emergence on new subclones and branching out of the tumor tissue, metastatic tissues and PDC as shown in Figure 13B. The order of mutation acquisition and subsequent branching of the samples reflects the clonal evolution and genetic drift observed in patient RCC.3. In this case the PDCs also reflect similar pattern of clonal evolution (except vena cava PDC), thereby indicating that the PDCs not just shared the subclonal mutations but could also recapitulate the clonal evolutionary processes.

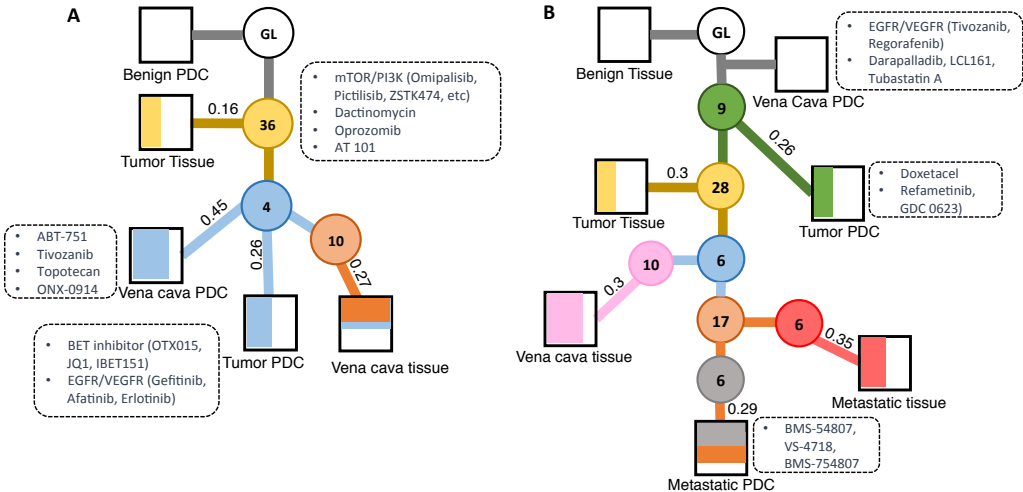


Figure 13: Genetic and drug response heterogeneity profiles of RCC.1 (A) and RCC.3 (B). Phylogenetic analysis of somatic variants from tumor, metastatic tissues and their respective PDCs revealed the cancer cell lineage in the patient cases from the germline cell (GL). Numbers indicated in the circle reflect the number of variants that form part of the subclone. The average VAF of the subclones are indicated on the branches. The square boxes represent the tumor sample with the composing subclones. Based on the drug response profiles of the PDCs, we identified drugs that were responding specifically in tumor or metastatic PDCs. Top differentially responding drugs are shown next to the corresponding PDC in the figure.

3.2 Impact of genomic heterogeneity on drug response profiles

In order to study the impact of genomic heterogeneity on drug responses, we integrated the phylogenetic classification of tumor subclones with drug

sensitivity and resistance data of individual PDCs from tumor and metastatic sites to >450 approved and investigational oncology drugs. We observed interesting correlations between site-specific mutations and response to targeted therapies. In patient case RCC.1, the ancestral clone carried mutations in *VHL*, *PIK3C2A*, *PBRM1*, *KDM5C* and copy number alterations in gene loci for *EGFR* and *CDKN2A/B* (Figure 13A). We observed significant response of mTOR/PI3K inhibitors such as omipalisib, pictilisib, AZD2014 in both primary and metastatic PDCs, indicating that the common shared mutations could underlie comparable but significant responses to mTOR/PI3K inhibitors, which are currently being investigated for treatment of RCC patients.¹²³ We observed selective responses in the primary and metastatic PDCs which can be hypothesized to be the result of site-specific subclones that were identified based on the phylogenetic analysis. Importantly, we observed significantly high response to *BET* inhibitors JQ1 and OTX015 and VEGFR inhibitors gefitinib and erlotinib in primary tumor PDC, which showed no response in the metastatic PDC. Conversely, we found anti-mitotic inhibitor ABT-751 and topoisomerase inhibitor topotecan (approved ovarian cancer drug¹²⁴) to be highly sensitive in metastatic PDC but not primary tumor PDC.

In patient case RCC.3 we observed significant response to mTOR/PI3K inhibitors (AZD2014, omipalisib and ZSRK474), *BET* inhibitors (JQ1, OTx015 and I-BET151) and topoisomerase inhibitors (camptothecin, topotecan and SN-38) across all three PDCs (Figure 13B). Presence of *BRD4* mutations in the ancestral clone shared among the PDCs could explain the response to *BET* inhibitors.¹²⁵⁻¹²⁷ Distinct PDC-specific responses were observed in RCC.3 case as well. Tumor and secondary metastatic PDCs which were genetically distant from vena cava metastatic PDC showed sensitivity to mTOR inhibitor temsirolimus, an approved drug for treating RCC.^{42,128} Interestingly, we noted VEGFR inhibitors tivozanib and regorafenib to be highly sensitive in vena cava metastatic PDC in comparison to the other two PDCs. Both these cases clearly illustrate the underlying intra-tumoral heterogeneity of not just genomic alterations but also drug responses in RCC patients.

DISCUSSION

The work presented in this thesis were designed to build on the FIMM individualized systems medicine (ISM) platform by studying how cancer evolves in time and space, while taking into consideration the impact of therapy. We aimed to develop methods that would allow the monitoring of the impact of treatment choices made and retrospectively understanding the mechanisms behind drug responses and drug resistance in heterogeneous, evolving tumors. The methods developed as part of the thesis are applicable to any cancer type, wherein longitudinal or spatial samples are acquired in a personalized medicine setting. Further, I will expound on the major lessons learnt from the studies in this section.

1. SUBCLONAL RESPONSES TO THERAPIES

As part of our individualized medicine strategy, we developed and implemented novel analytical tools to track subclonal responses following chemotherapy and tailored patient treatments.

In Study I, we tracked subclonal-responses to therapy in 13 AML patients from diagnosis to later relapse. Six patients received tailored treatments based on the ISM approach. In this study, we established methods to i) enhance rare variant calling by utilizing UMI-based ultra-deep resequencing approach, ii) quantify subclonal composition by implementing correlation and linear-trend analysis, and iii) track subclonal responses to therapy from diagnosis to relapse. Previous studies on clonal evolution with whole exome or genome sequencing at 100-300X coverage provide sensitivity to reliably track only clones present at greater than 5-10% frequency.^{106,129} Detection of low-frequency subclones by genome-wide sequencing is technically challenging and obviously not feasible at this depth.^{48,106,129} Ultra-deep amplicon resequencing therefore complements exome sequencing, providing a comprehensive picture of the clonal evolution in AML during disease progression by enabling detection of rare subclonal variants at specific sites of the genome. Using this approach, we achieved an average sequencing coverage of 68,500x, allowing the detection of subclones and possibly pre-leukemic subclones with frequency < 0.3% with frequency < 0.3%.

From the 13 AML cases, we observed significant differential responses of leukemic subclones to both chemo and targeted therapies. Of the six patients who received tailored targeted therapies, we observed significant subclonal responses in five patients. For example, patient case #560 was categorized as a non-responder with recurrent disease based on ELN criteria¹¹⁹ but we quantified significant response of all major subclones to temsirolimus-dasatinib therapy. Therefore, we believe that estimations of therapy responses utilizing subclone-level data provides a more comprehensive, quantitative and biologically more relevant picture of the molecular impact of the treatment. Drug response data on cell subclones will provide valuable insights on the cause of resistance. Therefore subsequent therapies can be targeted to inhibit the potentially emerging resistant subclones when possible. Further, we detected new subclones arising specifically at relapse and expansion of rare subclones at relapse that were pre-existing at diagnosis. These minor subclones were present at very high frequency at relapse, indicating therapy-based selection leading to selection of drug-resistant subclones. Taken together, we believe that monitoring of subclonal responses to therapy, enabled by ultra-deep amplicon sequencing approaches, provides valuable insights on drug resistance and sensitivities which forms an important basis for future drug development and personalized medicine strategies. In this study, we retrospectively confirmed that relapse-specific subclonal variants existed already at diagnosis as low frequency subclones. Therefore, it is critical to develop approaches that can identify these variants prospectively.

Overall, we saw that chemotherapy was effective as a first-line therapy but for relapse patients resistant to chemotherapeutics, understanding the behaviour of subclonal responses to targeted therapies will provide us insights on the impact of targeted therapies on targeted subclones which could be missed when response is measured based on clinical evaluation.

2. MECHANISTIC INSIGHTS ON GENETIC HETEROGENEITY UNDERLYING DRUG RESPONSE AND RESISTANCE

By tracking treatment responses of cancer subclones, we gained valuable insights on mechanisms underlying drug response or resistance. Of the six patients in Study I who received targeted therapies, we were able to hypothesize the underlying mechanism of drug resistance and sensitivity in five patients (Table 1). For example, response to dasatinib-sunitinib-temsirolimus therapy in patient #600 was mediated by *FLT3* mutations. Similarly, patient #1186 showed response to azacitidine-sunitinib combination therapy, sunitinib sensitivity was mediated by the presence of mutation in *PDGFRB* fusion and *FLT3*. Based on significant subclonal response of *DNMT3A*, *FLT3* and *NPM1* mutated subclones to therapy, we hypothesise these mutations were underlying azacitidine sensitivity. Mutations in *DNMT3A* have been reported to sensitize patients to hypomethylating agents.^{130,131} Intrinsic resistance of patient #3443 and acquired resistance in #784 could also be explained based on their subclonal profiles. In patient #3443 we identified loss of *NF1* gene at diagnosis and first relapse which explains the intrinsic resistance of this patient to cytarabine-based therapy.¹³²⁻¹³⁴ Subsequently, the patient relapsed with an additional loss of *NF1* gene copy and deletion in *DCK* gene, providing mechanistic explanations to the observed resistances towards cytarabine- and subsequent ruxolitinib-trametinib -based therapies.¹³⁴⁻¹³⁷

PATIENT #	TAILORED TREATMENT RECIEVED	CLINICAL RESPONSE	SUBCLONAL RESPONSE	MECHANISTIC INSIGHT
560	Temsirolimus-Dasatinib	Recurring disease	Sensitive	Not known
600	Dasatinib-Sunitinib-Temsirolimus	Complete Remission	Sensitive	Driven by FLT3 inhibition
784	Dasatinib-Sunitinib-Temsirolimus	Recurring disease	Acquired resistance from fusion genes	Resistance driven by selection of ETV6-NTRK3. Hyperactivation of mTORC1
1145	Ruxolitinib-dexamethasone	Recurring disease	Resistant	Mediated by TP53
1886	Azacitidine-Sunitinib	Complete Remission	Sensitive	Response driven by DNMT3A, FLT3 and NPM1 subclone and PDGFRB fusion.
3443	Ruxolitinib-trametinib	Recurring disease	Resistant	NF1 and DCK1 loss mediates resistance

Table 1: Mechanistic insights on drug-gene interactions

In Study III, by integrating drug response and subclonal information from PDCs we identified subgroups of drugs that were likely to work on specific subclones. We identified a subgroup of drugs that were responsive in all PDCs that correspond to early mutations shared among the primary and metastatic sites. PDCs with shared mutations showed distinct drug response profiles that correlated with the genetic heterogeneity. Further, site-specific drug responses were observed for the PDCs, which in some cases correlated with the unique mutations identified in the site. Overall, we observed intra-tumor heterogeneity not just in the genetic composition but also in drug responsiveness.

Taken together, we showed how tracking of subclonal architecture with treatment response could give mechanistic and pharmacogenomics insights on drug sensitivities and resistances. By inferring subclone-based responses, we can design intelligent therapeutic combinations that target multiple subclones, achieve longer remissions and prevent relapses or metastasis.

3. LESSONS LEARNT FROM PERSONALIZED TREATMENT OF PATIENTS

Personalizing patient treatment with the aim to achieve greater impact and survival benefits has its advantages and limitations. In our individualized medicine program, involving clinical translation of DSRT and genome sequencing data in AML and RCC, we have gained better understanding on why and how some therapies could fail. Here I present my perspective of the lessons learnt from personalizing cancer therapy for patients.

Temporal and Spatial tumor sampling: Acquiring and analyzing of multiple tumor samples from diagnosis to later relapses, metastasis and specifically before and after therapy forms the most important basis for implementing personalized therapy and monitoring its benefit. Taking tumor samples from multiple sites is necessary to gain a comprehensive estimation of intra-tumor heterogeneity while reducing sampling bias. We observed this effect to be more profound in solid tumors such as RCC than in AML, where the genetic heterogeneity between multiple sites was not as significant as in solid tumors. Secondly, given that clonal evolution of cancers is common, temporal sampling of the tumor should be part of any personalized medicine strategy. We showed in our studies how temporal sampling of tumor enabled us to monitor therapy response at subclone level and to gain understanding of the therapy responses, which differed from standard ways to measure clinical response in patients.

As discussed earlier, tumors evolve in response to therapy and by studying responses of subclones from samples taken before, after and during therapy we could identify the source of drug-resistant subclones seen at relapse, differential response of subclones to chemo and targeted therapies and mechanisms underlying drug sensitivities and resistance. Therefore, spatial and temporal sampling of tumor is necessary to understand the nature of cancer in an individual patient before designing personalized treatment options. Although it might not be clinically feasible in all cancers to obtain patient biopsies frequently, approaches such as liquid biopsies can be considered. However, the measurement of blood clonal composition in liquid or solid tumors represents a summary of

the responses at multiple sites, and it cannot reveal the same subclonal data as biopsies from individual sites.

Adapting and developing methodologies: Technological improvements in sequencing, new discoveries such as CRISPR-Cas9 genome editing technology and other biological advancements result in accumulation of new, large-scale data. In the context of personalized medicine, new insights on cancer biology, drug discoveries, and therapy regimens call for an improved method to integrate large-scale data to benefit a single patient. For example, we adapted an existing deep-sequencing approach to enhance rare variant calling. Similarly, new tools have been developed to identify subclonal composition, driver mutations, intra-tumoral heterogeneity and inferring cancer –cell lineage to expand our understanding of cancer biology while addressing the dynamic nature of evolution. With the ever-increasing amount of data that is generated, new analytical methods to optimally utilize the data and integrate multiple data sets are required.

Why do some patients respond to treatment while others do not? Clinical translation of research findings does not always yield the desired benefit. While translating our findings to clinical use in patient treatment, we saw examples of *ex vivo* responses correlating with patient response as well as frequent cases where *ex vivo* results did not translate *in vivo*. There are many factors that contribute to the failure of translating *ex vivo* responses *in vivo*. One of the major factors being the fact that cells in culture do not replicate the tumor microenvironment resulting in drug responses that cannot be translated clinically. In one such study, finding from our ISM platform showed how cells cultured in bone marrow stroma-derived conditional media showed significantly reduced sensitivity of BCL2 inhibitors and increased sensitivity to JAK inhibitors in comparison to standard cell culture media.¹³⁸ The study highlights the impact micro-environment can have on drug sensitivities and the importance of cell culture media conditions in mimicking *in vivo* conditions.

Another crucial factor that impacts the clinical translation of targeted therapies is the clonal and subclonal composition of the patient. Given the heterogeneity of cancer, it is critical to identify if the targeted genes are clonal or subclonal. If the

targeted gene is subclonal, the major clone present in the tumor is not eradicated. Also, as shown in our study and other previous reports^{106,107,139}, subclones that are in low frequency can emerge to dominate at relapse therefore underscoring the importance of targeting multiple subclones.

Retrospective versus prospective analysis: Time is a crucial factor for translating personalized molecular data to the clinic, especially in advanced, fast progressing cancers. In order to enable translation of drug response, genomic, epigenomic, proteomic and other related patient data, it is critical to generate, analyze and assimilate these data into meaningful clinical recommendations in a matter of days and at the level of $n=1$, without the ability to use classical statistical approaches. Therefore, in the majority of our cases clinical translation in chemo-refractory AML has been based on *ex vivo* drug response data that is generated and interpreted within 3-4 days. Current genomic sequencing analysis is not as rapid to be available along with drug response data, but the falling costs and improved sequencing technologies will enable faster generation and analysis of data in near future. Therefore, analysis of clonal evolution, heterogeneity and subclonal responses to therapy is currently done retrospectively rather than prospectively. Although this might not benefit the patient in question, we gain valuable insights on subclonal dynamics in response to therapy, correlation between subclone and drug responses, mechanistic insights on therapy that will undoubtedly benefit future patient care and drug discoveries. It is important to deliberate on strategies that can enable prospective analysis of subclonal emergence and therapy that will help us target clonal evolution and heterogeneity thereby preventing relapse or metastasis. Therefore, both retrospective and prospective analysis of clonal evolution and heterogeneity will have great impact on cancer treatment both at personalized and cohort level.

4. NEXT STEPS: TARGETING CLONAL HETEROGENEITY AND EVOLUTION

The first step in solving a complexity is recognizing and understanding the extent of it. Our understanding of clonal evolution and heterogeneity of cancers has rapidly increased in the last eight years with the application of massively parallel

sequencing divulging cancer evolution and heterogeneity from spatially and temporally acquired tumor samples. By studying the impact of treatment on subclones and vice versa, we can redefine our personalized therapy and drug development strategies to target clonal evolution and heterogeneity. We are now starting to piece together the puzzle of cancer. Recent interesting studies have shown how cancers can evolve in non-Darwinian fashion, wherein drug resistant cells are not developed based on selection pressure but in a Lamarckian fashion wherein the drug resistant phenotype is induced and enhanced by multiple processes.^{140,141}

Cancer is basically a composition of cells with extraordinary proliferating capacity that have evaded all known checkpoints to control the growth. In addition, it evolves dynamically by acquiring new mutations and adapts for survival. As stated by Mel Greaves, if we were to live long enough, we (all) will eventually develop “covert” cancer.¹⁴² In the light of evidence pointing to the extent of cancer heterogeneity and evolution underlying drug resistance, and the failure of both chemo and targeted drugs in achieving cancer cure, the question we need to ask is; *can cancer be cured?* Maybe this is not possible in all cases.

Therefore, many recent strategies rethinking the way to treat cancer are evolving. For instance, adaptive therapy proposed by Gatenby et al. advocates that eliminating all cancer cells eventually leads to emergence of resistant clones. Whereas, maintaining a stable cancer cell population causes less heterogeneity and competing evolution that could result in long-term survival. The approach essentially limits the heterogeneity of cancer subclones by eliminating selective pressures and maintaining drug responsive cell populations. The authors have shown evidence of adaptive therapy achieving long-term survival in ovarian cancer mice models injected with carboplatin therapy in multiple doses based on tumor size, wherein the dosage was reduced if the tumor reduced in size. The adaptive therapy strategy has its main advantage in addressing the evolutionary complexity of tumors. Although clinical adoption of this strategy might be challenging where appropriate estimation of treatment dosage versus tumor burden needs to be calculated to maintain “just enough” heterogeneity.

Similarly, other methods have been proposed to address tumor heterogeneity and evolution while treating cancer. Swanton et al. proposed the targeting of trunk antigens, which would result in higher causality of cancer cells rather than aiming for antigens targeting subclones that constitute the branches. One such approach involves identifying vulnerabilities acquired by cancer cells after developing resistance to drug A. Drug B that is sensitive in the new cancer might select for the pre-existing minor subclone that was response to drug A. The method basically involves exploiting vulnerabilities while balancing for resistance.

All these methods have their benefits and drawbacks, but the evolving treatment strategies will lead way to better treatment and management of cancer. Recent advances in immunotherapies for cancer, especially in leukemias are very promising. The approach of de-bulking the tumor initially with chemotherapy or targeted therapy and subsequently using immunotherapy to prevent future recurrence has its merit in combining existing treatment regimens with novel immunotherapies.^{143,144} The growing idea of maintaining cancer as a chronic disease while reducing cancer cell heterogeneity is another way forward. Further studies on the applicability and impact of this strategy are required. Overall, personalized medicine strategy that takes into consideration clonal evolution and heterogeneity will provide us valuable insights on treating cancer at both individual and population level.

“It is not our part to master all the tides of the world, but to do what is in us for the succour of those years wherein we are set, uprooting the evil in the fields that we know, so that those who live after may have clean earth to till.”

- Gandalf, The Lord of the Rings by J.R.R. Tolkien

ACKNOWLEDGEMENTS

This thesis was carried out at Institute for Molecular Medicine Finland (FIMM) as part of the Personalized Medicine Helsinki project. I would like to extend my thanks to everyone at the Institute for making it an amazing and inspiring place to work at.

I am grateful for the financial support received from the Doctoral Programme in Biomedicine (DPBM) , K. Albin Johansson Foundation and Maud Kuistila Memorial Foundation. I would especially like to thank the DPBM for the continued support during my PhD studies and for organizing excellent courses and events.

I would like to extend my heartfelt gratitude to my supervisors Olli Kallioniemi and Maija Wolf without whom this thesis would not have been possible. Olli, you have always been an inspiration for me and there are so many things I learnt from you. Your ability to look at the bigger picture in science as well as knowing the nitty-gritty details of things always amazed me. I want to thank you for giving me the freedom to explore my scientific horizons and for allowing me to pursue my extra-curricular activities during my studies.

Maija, no words will suffice my gratitude towards you for going above and beyond to help me complete my PhD. It wouldn't be an exaggeration to say that this thesis would not have been possible without your constant support, encouragement and help. You have inspired me both personally and professionally. Your non-compromising fairness and patience along with your amazing commitment to good science not just inspired me but also provided me with a role model in life. Thank you very much for being so amazing, I couldn't have asked for a better mentor in life! I would also like to thank Henrik Edgren, who was briefly my co-supervisor and played a great part in setting up my projects and find my footing in PhD.

I would like to thank my thesis committee members, Petri Bono, Dario Greco and Harri Lähdesmäki for their encouragement and commitment to help me with my thesis. Your scientific insights and comments on my thesis immensely motivated me during the course of my PhD. I am deeply grateful to Olli Lohi and

Merja Heinäniemi for reviewing my thesis and providing insightful comments that not just helped push my thesis up a notch, but also helped broaden my thinking of the field.

I have been fortunate to work with some brilliant minds with whom I co-authored the papers in this thesis. Mika Kontro, thank you very much for your valuable clinical insights and major contributions to the papers. Your honest efforts to help the patients you treat and beyond helped me play a small part of the narrative, I hope. Special shout out to Pekka Ellonen, Tiina Hanunnen, Sonja Langstrom, Maija Lepisto and Aino Palva from the sequencing team whose work formed a major part of my thesis. Many thanks to Timo Miettinen and Samuli Eldfors for your contributions. Henriikki Almusa, thank you so much for being such a great friend and contributing towards the publications. I really enjoyed our honest friendship and the game nights! Many thanks to Khalid Saeed and Paivi Ostling for bringing me onboard the RCC project, without which this thesis would be incomplete.

It has been a great experience working with the personalized medicine group. I would especially like to thank supervisors Caroline Heckman, Kimmo Porkka, Krister Wennerberg, Tero Aittokallio and Satu Mustjoki for their guidance and excellent collaboration. I learnt a lot from the PM level both at scientific and personal level during the last six years! I would like to extend my thanks to my fellow-PhD's with whom I shared the high's and low's of the process especially Disha Malani, Riikka Karjalainen, Tea Pemovska, Bhagwan Yadav and Mamun. I really enjoyed the AACR conferences in your company.

Without the support of my amazing lab-mates the last six years would have been extremely dull. Special thanks to Susanne, Khalid and Disha for our innumerable philosophical discussions and sharing the best and worst days together. Many thanks to my lab mates Astrid, Mariliina, Piia, Vilja, Sami, Vesa, Katja, Maija Puhka, Ruusu and Taija for the numerous laughs and times we spent together. Many thanks to Susanna Rosas for all the help.

I want to thank my colleagues at FIMM for the excellent friendships, warm and vibrant environment. I would like to thank Gretchen Repasky for a lot of things! To quote a few, thank you for all the help with FIMM-EMBL PhD training and encouraging me to organize multiple events at FIMM. I would like to thank my

lovely Think Different team, Christian, Alok, Meri, Prson and Anna for all the amazing work we did together. I would like to thank all my Indian friends Swapnil, Himanshu, Raghu, Ashwini Nagraj, Kiran, Ashwini Kumar, Sarang, Balaguru, Om and Rubina at FIMM for the fun get together's and lunch discussions. Thank you

Finally, I want to thank my amazing and loving family. For whatever I may have or lose in life, I treasure my family the most. Thank you amma and nanna (Lakshmi and Vasudeva Rao) for making our education a priority over anything else and your constant love and support throughout my life. Thank you akaji (Sravanthi) for being the best sister, mentor and friend in my life. I have feared and admired no one greater than thou! Thanks to my wonderful in-laws Kari and Liisa who have supported me through the PhD and beyond.

Most of all, I want to thank my husband Tuomas for being my rock in life. Without your encouragement, love, support and sacrifices I would never be able to realize my full potential. Thanks to my brat, Rian for keeping the joy alive.

Poojitha Ojamies

Helsinki 02.07.2018

REFERENCES

1. DeVita VT, Jr., Chu E. A history of cancer chemotherapy. *Cancer Res.* 2008;68(21):8643-8653.
2. Morgan G, Ward R, Barton M. The contribution of cytotoxic chemotherapy to 5-year survival in adult malignancies. *Clin Oncol (R Coll Radiol)*. 2004;16(8):549-560.
3. Miller KD, Siegel RL, Lin CC, et al. Cancer treatment and survivorship statistics, 2016. *CA Cancer J Clin.* 2016;66(4):271-289.
4. Davidoff AJ, Tang M, Seal B, Edelman MJ. Chemotherapy and survival benefit in elderly patients with advanced non-small-cell lung cancer. *J Clin Oncol.* 2010;28(13):2191-2197.
5. Rivera GK, Raimondi SC, Hancock ML, et al. Improved outcome in childhood acute lymphoblastic leukaemia with reinforced early treatment and rotational combination chemotherapy. *Lancet.* 1991;337(8733):61-66.
6. Rutkowski S, Bode U, Deinlein F, et al. Treatment of early childhood medulloblastoma by postoperative chemotherapy alone. *N Engl J Med.* 2005;352(10):978-986.
7. Burnett A, Wetzler M, Lowenberg B. Therapeutic advances in acute myeloid leukemia. *J Clin Oncol.* 2011;29(5):487-494.
8. Dohner H, Weisdorf DJ, Bloomfield CD. Acute Myeloid Leukemia. *N Engl J Med.* 2015;373(12):1136-1152.
9. Holohan C, Van Schaeybroeck S, Longley DB, Johnston PG. Cancer drug resistance: an evolving paradigm. *Nat Rev Cancer.* 2013;13(10):714-726.
10. Chabner BA, Roberts TG, Jr. Timeline: Chemotherapy and the war on cancer. *Nat Rev Cancer.* 2005;5(1):65-72.
11. Chin L, Andersen JN, Futreal PA. Cancer genomics: from discovery science to personalized medicine. *Nat Med.* 2011;17(3):297-303.
12. Sawyers C. Targeted cancer therapy. *Nature.* 2004;432(7015):294-297.
13. Gorre ME, Mohammed M, Ellwood K, et al. Clinical resistance to STI-571 cancer therapy caused by BCR-ABL gene mutation or amplification. *Science.* 2001;293(5531):876-880.
14. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Jr., Kinzler KW. Cancer genome landscapes. *Science.* 2013;339(6127):1546-1558.
15. Tamborero D, Gonzalez-Perez A, Perez-Llamas C, et al. Comprehensive identification of mutational cancer driver genes across 12 tumor types. *Sci Rep.* 2013;3:2650.

16. Gonzalez-Perez A, Mustonen V, Reva B, et al. Computational approaches to identify functional genetic variants in cancer genomes. *Nat Methods*. 2013;10(8):723-729.
17. Dees ND, Zhang Q, Kandoth C, et al. MuSiC: identifying mutational significance in cancer genomes. *Genome Res*. 2012;22(8):1589-1598.
18. Lawrence MS, Stojanov P, Polak P, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature*. 2013;499(7457):214-218.
19. Gonzalez-Perez A, Lopez-Bigas N. Functional impact bias reveals cancer drivers. *Nucleic Acids Res*. 2012;40(21):e169.
20. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res*. 2003;31(13):3812-3814.
21. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010;7(4):248-249.
22. Greenman C, Stephens P, Smith R, et al. Patterns of somatic mutation in human cancer genomes. *Nature*. 2007;446(7132):153-158.
23. Tamborero D, Gonzalez-Perez A, Lopez-Bigas N. OncodriveCLUST: exploiting the positional clustering of somatic mutations to identify cancer genes. *Bioinformatics*. 2013;29(18):2238-2244.
24. Papaemmanuil E, Gerstung M, Bullinger L, et al. Genomic Classification and Prognosis in Acute Myeloid Leukemia. *N Engl J Med*. 2016;374(23):2209-2221.
25. Cancer Genome Atlas Research N, Ley TJ, Miller C, et al. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med*. 2013;368(22):2059-2074.
26. Corces-Zimmerman MR, Hong WJ, Weissman IL, Medeiros BC, Majeti R. Preleukemic mutations in human acute myeloid leukemia affect epigenetic regulators and persist in remission. *Proc Natl Acad Sci U S A*. 2014;111(7):2548-2553.
27. Vardiman JW, Thiele J, Arber DA, et al. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood*. 2009;114(5):937-951.
28. McKerrell T, Park N, Moreno T, et al. Leukemia-associated somatic mutations drive distinct patterns of age-related clonal hemopoiesis. *Cell Rep*. 2015;10(8):1239-1245.
29. Jaiswal S, Fontanillas P, Flannick J, et al. Age-related clonal hematopoiesis associated with adverse outcomes. *N Engl J Med*. 2014;371(26):2488-2498.

30. Young AL, Challen GA, Birmann BM, Druley TE. Clonal haematopoiesis harbouring AML-associated mutations is ubiquitous in healthy adults. *Nat Commun*. 2016;7:12484.
31. Hsieh JJ, Purdue MP, Signoretti S, et al. Renal cell carcinoma. *Nat Rev Dis Primers*. 2017;3:17009.
32. Cancer Genome Atlas Research N. Comprehensive molecular characterization of clear cell renal cell carcinoma. *Nature*. 2013;499(7456):43-49.
33. Sato Y, Yoshizato T, Shiraishi Y, et al. Integrated molecular analysis of clear-cell renal cell carcinoma. *Nat Genet*. 2013;45(8):860-867.
34. Dalglish GL, Furge K, Greenman C, et al. Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. *Nature*. 2010;463(7279):360-363.
35. Varela I, Tarpey P, Raine K, et al. Exome sequencing identifies frequent mutation of the SWI/SNF complex gene PBRM1 in renal carcinoma. *Nature*. 2011;469(7331):539-542.
36. Kibel A, Iliopoulos O, DeCaprio JA, Kaelin WG, Jr. Binding of the von Hippel-Lindau tumor suppressor protein to Elongin B and C. *Science*. 1995;269(5229):1444-1446.
37. Rathmell WK, Chen S. VHL inactivation in renal cell carcinoma: implications for diagnosis, prognosis and treatment. *Expert Rev Anticancer Ther*. 2008;8(1):63-73.
38. Maxwell PH, Wiesener MS, Chang GW, et al. The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis. *Nature*. 1999;399(6733):271-275.
39. Massari F, Ciccicarese C, Santoni M, et al. Metabolic alterations in renal cell carcinoma. *Cancer Treat Rev*. 2015;41(9):767-776.
40. Robb VA, Karbowiczek M, Klein-Szanto AJ, Henske EP. Activation of the mTOR signaling pathway in renal clear cell carcinoma. *J Urol*. 2007;177(1):346-352.
41. Motzer RJ, Escudier B, Oudard S, et al. Efficacy of everolimus in advanced renal cell carcinoma: a double-blind, randomised, placebo-controlled phase III trial. *Lancet*. 2008;372(9637):449-456.
42. Hudes G, Carducci M, Tomeczak P, et al. Temsirolimus, interferon alfa, or both for advanced renal-cell carcinoma. *N Engl J Med*. 2007;356(22):2271-2281.
43. Gerlinger M, Horswell S, Larkin J, et al. Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. *Nat Genet*. 2014;46(3):225-233.

44. Gerlinger M, Rowan AJ, Horswell S, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med.* 2012;366(10):883-892.
45. Nowell PC. The clonal evolution of tumor cell populations. *Science.* 1976;194(4260):23-28.
46. Anderson K, Lutz C, van Delft FW, et al. Genetic variegation of clonal architecture and propagating cells in leukaemia. *Nature.* 2011;469(7330):356-361.
47. Campbell PJ, Yachida S, Mudie LJ, et al. The patterns and dynamics of genomic instability in metastatic pancreatic cancer. *Nature.* 2010;467(7319):1109-1113.
48. Ding L, Ley TJ, Larson DE, et al. Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature.* 2012;481(7382):506-510.
49. Walter MJ, Shen D, Ding L, et al. Clonal architecture of secondary acute myeloid leukemia. *N Engl J Med.* 2012;366(12):1090-1098.
50. Wu X, Northcott PA, Dubuc A, et al. Clonal selection drives genetic divergence of metastatic medulloblastoma. *Nature.* 2012;482(7386):529-533.
51. Yachida S, Jones S, Bozic I, et al. Distant metastasis occurs late during the genetic evolution of pancreatic cancer. *Nature.* 2010;467(7319):1114-U1126.
52. Welch JS, Ley TJ, Link DC, et al. The origin and evolution of mutations in acute myeloid leukemia. *Cell.* 2012;150(2):264-278.
53. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature.* 2001;409(6822):860-921.
54. Hayden EC. Technology: The \$1,000 genome. *Nature.* 2014;507(7492):294-295.
55. Darwin C. *On the origin of species by means of natural selection.* London,: J. Murray; 1859.
56. Quail DF, Joyce JA. Microenvironmental regulation of tumor progression and metastasis. *Nat Med.* 2013;19(11):1423-1437.
57. Joyce JA, Pollard JW. Microenvironmental regulation of metastasis. *Nat Rev Cancer.* 2009;9(4):239-252.
58. Bissell MJ, Hines WC. Why don't we get more cancer? A proposed role of the microenvironment in restraining cancer progression. *Nat Med.* 2011;17(3):320-329.
59. Tredan O, Galmarini CM, Patel K, Tannock IF. Drug resistance and the solid tumor microenvironment. *J Natl Cancer Inst.* 2007;99(19):1441-1454.

60. Alizadeh AA, Aranda V, Bardelli A, et al. Toward understanding and exploiting tumor heterogeneity. *Nat Med*. 2015;21(8):846-853.
61. Forsberg LA, Rasi C, Razzaghian HR, et al. Age-related somatic structural changes in the nuclear genome of human blood cells. *Am J Hum Genet*. 2012;90(2):217-228.
62. Lynch M. Evolution of the mutation rate. *Trends Genet*. 2010;26(8):345-352.
63. Lawrence MS, Stojanov P, Mermel CH, et al. Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature*. 2014;505(7484):495-501.
64. Taylor BS, Schultz N, Hieronymus H, et al. Integrative genomic profiling of human prostate cancer. *Cancer Cell*. 2010;18(1):11-22.
65. Berger MF, Lawrence MS, Demichelis F, et al. The genomic complexity of primary human prostate cancer. *Nature*. 2011;470(7333):214-220.
66. Cancer Genome Atlas Research N. Integrated genomic analyses of ovarian carcinoma. *Nature*. 2011;474(7353):609-615.
67. Greaves M, Maley CC. Clonal evolution in cancer. *Nature*. 2012;481(7381):306-313.
68. Jan M, Snyder TM, Corces-Zimmerman MR, et al. Clonal evolution of preleukemic hematopoietic stem cells precedes human acute myeloid leukemia. *Sci Transl Med*. 2012;4(149):149ra118.
69. Yates LR, Campbell PJ. Evolution of the cancer genome. *Nat Rev Genet*. 2012;13(11):795-806.
70. Nik-Zainal S, Van Loo P, Wedge DC, et al. The life history of 21 breast cancers. *Cell*. 2012;149(5):994-1007.
71. Gundem G, Van Loo P, Kremeyer B, et al. The evolutionary history of lethal metastatic prostate cancer. *Nature*. 2015;520(7547):353-357.
72. Navin N, Kendall J, Troge J, et al. Tumour evolution inferred by single-cell sequencing. *Nature*. 2011;472(7341):90-94.
73. de Bruin EC, McGranahan N, Mitter R, et al. Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. *Science*. 2014;346(6206):251-256.
74. Miyamoto T, Weissman IL, Akashi K. AML1/ETO-expressing nonleukemic stem cells in acute myelogenous leukemia with 8;21 chromosomal translocation. *Proc Natl Acad Sci USA*. 2000;97(13):7521-7526.
75. Shlush LI, Zandi S, Mitchell A, et al. Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. *Nature*. 2014;506(7488):328-333.
76. Baca SC, Prandi D, Lawrence MS, et al. Punctuated evolution of prostate cancer genomes. *Cell*. 2013;153(3):666-677.

77. Kloosterman WP, Hoogstraat M, Paling O, et al. Chromothripsis is a common mechanism driving genomic rearrangements in primary and metastatic colorectal cancer. *Genome Biol.* 2011;12(10):R103.
78. Gerlinger M, McGranahan N, Dewhurst SM, Burrell RA, Tomlinson I, Swanton C. Cancer: evolution within a lifetime. *Annu Rev Genet.* 2014;48:215-236.
79. Davis A, Gao R, Navin N. Tumor evolution: Linear, branching, neutral or punctuated? *Biochim Biophys Acta.* 2017;1867(2):151-161.
80. Sottoriva A, Kang H, Ma Z, et al. A Big Bang model of human colorectal tumor growth. *Nat Genet.* 2015;47(3):209-216.
81. Leibowitz ML, Zhang CZ, Pellman D. Chromothripsis: A New Mechanism for Rapid Karyotype Evolution. *Annu Rev Genet.* 2015;49:183-211.
82. Liu P, Erez A, Nagamani SC, et al. Chromosome catastrophes involve replication mechanisms generating complex genomic rearrangements. *Cell.* 2011;146(6):889-903.
83. Welch DR. Tumor Heterogeneity--A 'Contemporary Concept' Founded on Historical Insights and Predictions. *Cancer Res.* 2016;76(1):4-6.
84. Swanton C. Intratumor Heterogeneity: Evolution through Space and Time. *Cancer Research.* 2012;72(19):4875-4882.
85. Fidler IJ. Tumor heterogeneity and the biology of cancer invasion and metastasis. *Cancer Res.* 1978;38(9):2651-2660.
86. Kuukasjarvi T, Karhu R, Tanner M, et al. Genetic heterogeneity and clonal evolution underlying development of asynchronous metastasis in human breast cancer. *Cancer Res.* 1997;57(8):1597-1604.
87. Wolman SR. Cytogenetic heterogeneity: its role in tumor evolution. *Cancer Genet Cytogenet.* 1986;19(1-2):129-140.
88. Gorunova L, Hoglund M, Andren-Sandberg A, et al. Cytogenetic analysis of pancreatic carcinomas: intratumor heterogeneity and nonrandom pattern of chromosome aberrations. *Genes Chromosomes Cancer.* 1998;23(2):81-99.
89. Navin N, Krasnitz A, Rodgers L, et al. Inferring tumor progression from genomic heterogeneity. *Genome Research.* 2010;20(1):68-80.
90. Nik-Zainal S, Alexandrov LB, Wedge DC, et al. Mutational Processes Molding the Genomes of 21 Breast Cancers. *Cell.* 2012;149(5):979-993.
91. McGranahan N, Swanton C. Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. *Cell.* 2017;168(4):613-628.
92. Ojha J, Ayres J, Secreto C, et al. Deep sequencing identifies genetic heterogeneity and recurrent convergent evolution in chronic lymphocytic leukemia. *Blood.* 2015;125(3):492-498.

93. Miller KD, Siegel RL, Lin CC, et al. Cancer treatment and survivorship statistics, 2016. *Ca-Cancer J Clin.* 2016;66(4):271-289.
94. Parsons DW, Jones S, Zhang XS, et al. An integrated genomic analysis of human glioblastoma Multiforme. *Science.* 2008;321(5897):1807-1812.
95. Egan JB, Shi CX, Tembe W, et al. Whole-genome sequencing of multiple myeloma from diagnosis to plasma cell leukemia reveals genomic initiating events, evolution, and clonal tides. *Blood.* 2012;120(5):1060-1066.
96. Weston-Bell N, Gibson J, John M, et al. Exome sequencing in tracking clonal evolution in multiple myeloma following therapy. *Leukemia.* 2013;27(5):1188-1191.
97. Green MR, Gentles AJ, Nair RV, et al. Hierarchy in somatic mutations arising during genomic evolution and progression of follicular lymphoma. *Blood.* 2013;121(9):1604-1611.
98. Hunter C, Smith R, Cahill DP, et al. A hypermutation phenotype and somatic MSH6 mutations in recurrent human malignant gliomas after alkylator chemotherapy. *Cancer Research.* 2006;66(8):3987-3991.
99. Johnson BE, Mazor T, Hong CB, et al. Mutational Analysis Reveals the Origin and Therapy-Driven Evolution of Recurrent Glioma. *Science.* 2014;343(6167):189-193.
100. Yip S, Miao JY, Cahill DP, et al. MSH6 Mutations Arise in Glioblastomas during Temozolomide Therapy and Mediate Temozolomide Resistance. *Clin Cancer Res.* 2009;15(14):4622-4629.
101. Turke AB, Zejnullahu K, Wu YL, et al. Preexistence and Clonal Selection of MET Amplification in EGFR Mutant NSCLC. *Cancer Cell.* 2010;17(1):77-88.
102. Tzoneva G, Perez-Garcia A, Carpenter Z, et al. Activating mutations in the NT5C2 nucleotidase gene drive chemotherapy resistance in relapsed ALL. *Nature Medicine.* 2013;19(3):368-371.
103. Roche-Lestienne C, Soenen-Cornu V, Grardel-Duflos N, et al. Several types of mutations of the Abl gene can be found in chronic myeloid leukemia patients resistant to ST157, and they can pre-exist to the onset of treatment. *Blood.* 2002;100(3):1014-1018.
104. Shah NP, Nicoll JM, Nagar B, et al. Multiple BCR-ABL kinase domain mutations confer polyclonal resistance to the tyrosine kinase inhibitor imatinib (STI571) in chronic phase and blast crisis chronic myeloid leukemia. *Cancer Cell.* 2002;2(2):117-125.
105. Edwards SL, Brough R, Lord CJ, et al. Resistance to therapy caused by intragenic deletion in BRCA2. *Nature.* 2008;451(7182):1111-U1118.

106. Landau DA, Carter SL, Stojanov P, et al. Evolution and Impact of Subclonal Mutations in Chronic Lymphocytic Leukemia. *Cell*. 2013;152(4):714-726.
107. Ojames P, Kontro M, Edgren H, et al. Monitoring therapy responses at the leukemic subclone level by ultra-deep amplicon resequencing in acute myeloid leukemia. *Leukemia*. 2017;31(5):1048-1058.
108. Esposito A, Criscitiello C, Locatelli M, Milano M, Curigliano G. Liquid biopsies for solid tumors: Understanding tumor heterogeneity and real time monitoring of early resistance to targeted therapies. *Pharmacol Therapeut*. 2016;157:120-124.
109. Dawson SJ, Tsui DWY, Murtaza M, et al. Analysis of Circulating Tumor DNA to Monitor Metastatic Breast Cancer. *New Engl J Med*. 2013;368(13):1199-1209.
110. De Mattos-Arruda L, Weigelt B, Cortes J, et al. Capturing intra-tumor genetic heterogeneity by de novo mutation profiling of circulating cell-free tumor DNA: a proof-of-principle. *Ann Oncol*. 2014;25(9):1729-1735.
111. Murtaza M, Dawson SJ, Tsui DWY, et al. Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. *Nature*. 2013;497(7447):108-112.
112. Koboldt DC, Zhang Q, Larson DE, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res*. 2012;22(3):568-576.
113. Cingolani P, Platts A, Wang le L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)*. 2012;6(2):80-92.
114. Rubio-Perez C, Tamborero D, Schroeder MP, et al. In silico prescription of anticancer drugs to cohorts of 28 tumor types reveals targeting opportunities. *Cancer Cell*. 2015;27(3):382-396.
115. Olshen AB, Bengtsson H, Neuvial P, Spellman PT, Olshen RA, Seshan VE. Parent-specific copy number in paired tumor-normal studies using circular binary segmentation. *Bioinformatics*. 2011;27(15):2038-2046.
116. Pemovska T, Kontro M, Yadav B, et al. Individualized systems medicine strategy to tailor treatments for patients with chemorefractory acute myeloid leukemia. *Cancer discovery*. 2013;3(12):1416-1429.

117. Yadav B, Pemovska T, Szwajda A, et al. Quantitative scoring of differential drug sensitivity for individually optimized anticancer therapies. *Sci Rep.* 2014;4:5193.
118. Popic V, Salari R, Hajirasouliha I, Kashef-Haghighi D, West RB, Batzoglou S. Fast and scalable inference of multi-sample cancer lineages. *Genome Biol.* 2015;16:91.
119. Dohner H, Estey EH, Amadori S, et al. Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet. *Blood.* 2010;115(3):453-474.
120. Faivre S, Demetri G, Sargent W, Raymond E. Molecular basis for sunitinib efficacy and future clinical development. *Nature reviews Drug discovery.* 2007;6(9):734-745.
121. Fiedler W, Serve H, Dohner H, et al. A phase 1 study of SU11248 in the treatment of patients with refractory or resistant acute myeloid leukemia (AML) or not amenable to conventional therapy for the disease. *Blood.* 2005;105(3):986-993.
122. Yee KW, Schittenhelm M, O'Farrell AM, et al. Synergistic effect of SU11248 with cytarabine or daunorubicin on FLT3 ITD-positive leukemic cells. *Blood.* 2004;104(13):4202-4209.
123. Munster P, Aggarwal R, Hong D, et al. First-in-Human Phase I Study of GSK2126458, an Oral Pan-Class I Phosphatidylinositol-3-Kinase Inhibitor, in Patients with Advanced Solid Tumor Malignancies. *Clin Cancer Res.* 2016;22(8):1932-1939.
124. Bookman MA, Malmstrom H, Bolis G, et al. Topotecan for the treatment of advanced epithelial ovarian cancer: an open-label phase II study in patients treated after prior chemotherapy that contained cisplatin or carboplatin and paclitaxel. *J Clin Oncol.* 1998;16(10):3345-3352.
125. Filippakopoulos P, Qi J, Picaud S, et al. Selective inhibition of BET bromodomains. *Nature.* 2010;468(7327):1067-1073.
126. Mertz JA, Conery AR, Bryant BM, et al. Targeting MYC dependence in cancer by inhibiting BET bromodomains. *Proc Natl Acad Sci U S A.* 2011;108(40):16669-16674.
127. Delmore JE, Issa GC, Lemieux ME, et al. BET bromodomain inhibition as a therapeutic strategy to target c-Myc. *Cell.* 2011;146(6):904-917.

128. Motzer RJ, Hudes GR, Curti BD, et al. Phase I/II trial of temsirolimus combined with interferon alfa for advanced renal cell carcinoma. *J Clin Oncol*. 2007;25(25):3958-3964.
129. Landau DA, Carter SL, Getz G, Wu CJ. Clonal evolution in hematological malignancies and therapeutic implications. *Leukemia*. 2014;28(1):34-43.
130. Metzeler KH, Walker A, Geyer S, et al. DNMT3A mutations and response to the hypomethylating agent decitabine in acute myeloid leukemia. *Leukemia*. 2012;26(5):1106-1107.
131. Traina F, Visconte V, Elson P, et al. Impact of molecular mutations on treatment response to DNMT inhibitors in myelodysplasia and related neoplasms. *Leukemia*. 2014;28(1):78-87.
132. Yin B, Morgan K, Hasz DE, Mao Z, Largaespada DA. Nf1 gene inactivation in acute myeloid leukemia cells confers cytarabine resistance through MAPK and mTOR pathways. *Leukemia*. 2006;20(1):151-154.
133. Cai J, Damaraju VL, Groulx N, et al. Two distinct molecular mechanisms underlying cytarabine resistance in human leukemic cells. *Cancer Research*. 2008;68(7):2349-2357.
134. Galmarini CM, Thomas X, Calvo F, et al. In vivo mechanisms of resistance to cytarabine in acute myeloid leukaemia. *Brit J Haematol*. 2002;117(4):860-868.
135. Lauchle JO, Kim D, Le DT, et al. Response and resistance to MEK inhibition in leukaemias initiated by hyperactive Ras. *Nature*. 2009;461(7262):411-U110.
136. Brundage ME, Tandon P, Eaves DW, et al. MAF mediates crosstalk between Ras-MAPK and mTOR signaling in NF1. *Oncogene*. 2014;33(49):5626-5636.
137. Whittaker SR, Theurillat JP, Van Allen E, et al. A Genome-Scale RNA Interference Screen Implicates NF1 Loss in Resistance to RAF Inhibition. *Cancer discovery*. 2013;3(3):350-362.
138. Karjalainen R, Pemovska T, Popa M, et al. JAK1/2 and BCL2 inhibitors synergize to counteract bone marrow stromal cell-induced protection of AML. *Blood*. 2017;130(6):789-802.
139. Walter MJ, Shen D, Ding L, et al. Clonal Architecture of Secondary Acute Myeloid Leukemia. *New Engl J Med*. 2012;366(12):1090-1098.

140. Li S, Garrett-Bakelman FE, Chung SS, et al. Distinct evolution and dynamics of epigenetic and genetic heterogeneity in acute myeloid leukemia. *Nat Med*. 2016;22(7):792-799.
141. Pisco AO, Brock A, Zhou J, et al. Non-Darwinian dynamics in therapy-induced cancer drug resistance. *Nat Commun*. 2013;4:2467.
142. Greaves M. Does everyone develop covert cancer? *Nature Reviews Cancer*. 2014;14(4):209-210.
143. Barrett AJ, Le Blanc K. Immunotherapy prospects for acute myeloid leukaemia. *Clin Exp Immunol*. 2010;161(2):223-232.
144. Schmidt C. The benefits of immunotherapy combinations. *Nature*. 2017;552(7685):S67-S69.